

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 15:57:10 ; Search time 70.63 Seconds
(without alignments)
55.779 Million cell updates/sec

Title: US-09-641-104A-6
Perfect score: 204
Sequence: 1 HREGLLAIFKSGIPALVKM.....GSPVDSVLFYAITTLHLL 41
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	204	100.0	781	2 S35099	beta-catenin - Afr
2	204	100.0	781	2 A38973	beta-catenin - hum
3	204	100.0	781	2 S35091	beta-catenin - mou
4	191	93.6	621	2 S35092	plakoglobin - mou
5	191	93.6	744	2 A32905	plakoglobin, desmo
6	188	92.2	738	2 S35093	plakoglobin - Afri
7	188	92.2	843	2 T12689	armadillo segment
8	184	90.2	817	2 S33793	hypothetical prote
9	179	87.7	820	2 S33794	hypothetical prote
10	160	78.4	806	2 JC4835	beta-catenin - Hyd
11	104	51.0	678	2 T23341	beta-catenin - Cae
12	65	31.9	538	2 A86157	probable importin
13	64	31.4	531	2 T01516	SRP1 protein homol
14	64	31.4	531	2 T52099	probable nuclear t
15	61	29.9	660	2 T45588	arm repeat contain
16	59	28.9	468	2 AF2163	hypothetical prote
17	59	28.9	639	2 T06735	hypothetical prote
18	59	28.9	811	2 T43175	armadillo protein
19	58	28.4	578	2 S50446	VAC8 protein - yea
20	58	28.4	674	2 T48204	hypothetical prote
21	57	27.9	528	2 T52101	probable nuclear t
22	57	27.9	538	2 F86225	hypothetical prote
23	56	27.5	425	2 E83023	3-deoxy-D-manno-oc
24	56	27.5	465	2 S34929	probable membrane
25	56	27.5	527	2 T04329	importin alpha - t
26	55	27.0	402	2 T06629	hypothetical prote
27	55	27.0	447	2 AD1432	conserved hypothet
28	55	27.0	447	2 AD1433	conserved hypothet
29	55	27.0	571	2 B86150	hypothetical prote

30 54 26.5 423 2 T06774
31 54 26.5 423 2 T40294
32 53.5 26.2 458 2 H82200
33 53.5 26.2 911 2 S28498
34 53 26.0 281 2 F97844
35 53 26.0 323 2 E71180
36 53 26.0 1050 2 T26395
37 53 26.0 1247 2 T42209
38 52.5 25.7 311 2 B70083
39 52.5 25.7 351 2 E69218
40 52 25.5 372 2 I58141
41 52 25.5 387 2 D72638
42 52 25.5 521 2 JC5505
43 52 25.5 529 2 A56516
44 52 25.5 529 2 S57873
45 52 25.5 529 2 S57345

ALIGNMENTS

RESULT 1:

S35099 ...
beta-catenin - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C>Date: 10-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 21-Jul-2000
C:Accession: S35099
R:McCreary, P.D.; Turck, C.W.; Gumbiner, B.
Science 254, 1359-1361, 1991

A:Title: A homolog of the armadillo protein in Drosophila (plakoglobin) associated w
A:Reference number: S35099; MUID:92073903
A:Accession: S35099
A:Molecule type: mRNA
A:Residues: 1-781 <MCC>

A:Cross-references: GB:M77013; MID:g214020; PIDN:AAA49670.1; PID:g214021
C:Keywords: cytoskeleton

Query Match 100.0%; Score 204; DB 2; Length 781;
Best Local Similarity 100.0%; Pred. No. 2.2e-19;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HREGLLAIFKSGIPALVKMLGSPVDSVLFYAITTLHLL 41
|||||
Db 224 HREGLLAIFKSGIPALVKMLGSPVDSVLFYAITTLHLL 264

RESULT 2:

A38973
beta-catenin - human

C:Species: Homo sapiens (man)

C>Date: 26-Jan-1996 #sequence_revision 08-Feb-1996 #text_change 08-Oct-1999

C:Accession: A38973; S55356; S31988

R:Huelsken, J.; Birchmeier, W.; Behrens, J.

J. Cell Biol. 127, 2061-2069, 1994

A:Title: E-cadherin and APC compete for the interaction with beta-catenin and the cyt

A:Reference number: A38973; MUID:95105247

A:Accession: A38973

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-781 <HUE>

A:Cross-references: GB:Z19054; MID:g38519; PIDN:CAA79497.1; PID:g38520

A:Experimental source: placenta

R:Nollet, F.; Berx, G.; Molemans, F.; van Roy, F.

submitted to the EMBL Data Library, June 1995

A:Description: H. sapiens beta-catenin mRNA.

A:Reference number: S55356

A:Accession: S55356

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-781 <NOL>

A:Cross-references: EMBL:X87838; MID:g1154853; PIDN:CAA61107.1; PID:g860988

C:Comment: The cytoplasmic protein beta-catenin normally binds to a region in the cyt

microfilament network.
C:Comment: Cellular levels of beta-catenin are regulated in part by the adenomatous polyd cancerous cell growth.

C:Genetics:
A:Gene: GDB:CTNNB1; CTNNB
A:Cross-references: GDB:141922; OMIM:116806
A:Map position: 3p22-3p21.3
C:Keywords: apoptosis; carcinogenesis; cell adhesion; cytosol
F:151-676/Region: 40-residue repeats

Query Match 100.0%; Score 204; DB 2; Length 781;
Best Local Similarity 100.0%; Pred. No. 2.2e-19;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 41
|||||
Db 224 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 264

RESULT 3

S35091
beta-catenin - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999
C:Accession: S35091
R:Butz, S.; Stappert, J.; Weissig, H.; Kemler, R.
Science 257, 1142-1144, 1992
A:Title: Plakoglobin and beta-catenin: distinct but closely related.
A:Reference number: S35091; MUID:92376536
A:Accession: S35091
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-781 <BUT>
A:Cross-references: EMBL:M90364; NID:g192141; PIDN:AAA37280.1; PID:g192142
C:Keywords: cytoskeleton

Query Match 100.0%; Score 204; DB 2; Length 781;
Best Local Similarity 100.0%; Pred. No. 2.2e-19;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 41
|||||
Db 224 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 264

RESULT 4

S35092
plakoglobin - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 17-Mar-1999
C:Accession: S35092
R:Butz, S.; Stappert, J.; Weissig, H.; Kemler, R.
Science 257, 1142-1144, 1992
A:Title: Plakoglobin and beta-catenin: distinct but closely related.
A:Reference number: S35091; MUID:92376536
A:Accession: S35092
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-621 <BUT>
A:Cross-references: EMBL:M90365
C:Keywords: cytoskeleton

Query Match 93.6%; Score 191; DB 2; Length 621;
Best Local Similarity 92.7%; Pred. No. 9.8e-18;
Matches 38; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 41
|||||
Db 91 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 131

RESULT 5
A32905
plakoglobin, desmosomal - human
C:Species: Homo sapiens (man)
C:Date: 22-Nov-1989 #sequence_revision 22-Nov-1989 #text_change 29-Aug-1997
C:Accession: A32905
R:Frankel, W.W.; Goldschmidt, M.D.; Zimbelmann, R.; Mueller, H.M.; Schiller, D.L.; Cow
Proc. Natl. Acad. Sci. U.S.A. 86, 4027-4031, 1989
A:Title: Molecular cloning and amino acid sequence of human plakoglobin, the common j
A:Reference number: A32905; MUID:89264555
A:Accession: A32905
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-744 <FRA>
A:Cross-references: GB:M23410
C:Genetics:
A:Gene: GDB:JUP
A:Cross-references: GDB:126565; OMIM:173325
A:Map position: 7pter-7qter
C:Keywords: cytoskeleton

Query Match 93.6%; Score 191; DB 2; Length 744;
Best Local Similarity 92.7%; Pred. No. 1.2e-17;
Matches 38; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 41
|||||
Db 215 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 255

RESULT 6

S35093
plakoglobin - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jul-2000
C:Accession: S35093; S24636
R:Fouquet, B.; Zimbelmann, R.; Franke, W.W.
Differentiation 51, 187-194, 1992
A:Title: Identification of plakoglobin in oocytes and early embryos of Xenopus laevis
A:Reference number: S35093; MUID:93053332
A:Accession: S35093
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-738 <FOU>
A:Cross-references: EMBL:M95593; NID:g214656; PIDN:AAA49931.1; PID:g214657
R:Demarais, A.A. EMBL Data Library, June 1992
submitted to the EMBL Data Library, June 1992
A:Reference number: S24636
A:Accession: S24636
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 133-184, 'V', 186-225, 'T', 227-292 <DEM>
A:Cross-references: EMBL:X67078; NID:965252; PID:965253
C:Keywords: cytoskeleton

Query Match 92.2%; Score 188; DB 2; Length 738;
Best Local Similarity 90.2%; Pred. No. 3e-17;
Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 41
|||||
Db 211 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 251

RESULT 7

T12689
armadillo segment polarity protein - fruit fly (Drosophila melanogaster)
N:Alternate names: protein 8684.6
C:Species: Drosophila melanogaster
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 21-Jul-2000

```

Query Match          90.2%; Score 184; DB 2; Length 817;
Best Local Similarity 87.8%; Pred. NO. 1.2e-16;
Matches 36; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      1 HREGLLAIFKSGGIPALVKMGSPVDSVLFYAITTLHNLLL 41
      ||:||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      247 HQGLLTIFKSGGIPALVKLLSPVESVLFYAITTLHNLLL 287

RESULT      9
S33794
hypothetical protein - sea urchin (Tripneustes gratilla)
C:Species: Tripneustes gratilla
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Aug-1997
C:Accession: S33794
R:Rosenthal, E.
Biochim. Biophys. Acta 1173, 337-341, 1993
A:Title: Identification of homologues to beta-catenin/plakoglobin/armadillo in two inver
A:Reference number: S33793; MUID:93305730
A:Accession: S33794
A>Status: preliminary

```

```

Query Match          78.4%  Score 160;  DB 2;  Length 806;
Best Local Similarity 73.2%  Pred. No. 2e-13;
Matches 30;  Conservative 7;  Mismatches 4;  Indels 0;  Gaps 0;

QY  1  HREGLLAIFKSGGIPALVKMLGSPDVSVLFYATTLHNLL 41
      |||:|||||  |||:|||||:  |||:|||||:|||||
DB  263  HROGLMAIFKCGIPALVKLLGHRIEAVVFYATTLHNLL 323

RESULT 11
T23341
beta-catenin - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
C:Accession: T23341; T42221
R:Harris, B.
Submitted to the EMBL Data Library, November 1996
A:Reference number: Z19729
A:Accession: T23341
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-678 <NL>
A:Cross-references: EMBL:Z81564; PIDN:CA804572.1; GSPDB:GN00019; CESP:K05C4.6
A:Experimental source: clone K05C4
R:Costa, M.; Raich, W.; Agbunag, C.; Hardin, J.; Priess, J.R.
Submitted to the EMBL Data Library, July 1997
A:Description: A putative catenin-cadherin system mediates morphogenesis of the C. el
A:Reference number: Z22085
A:Accession: T42221
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-678 <COS>
A:Cross-references: EMBL:AF016853; PIDN:AAB94552.1
C:Genetics:
A:Gene: hmp-2; CESP:K05C4.6
A:Map position: 1
A:Introns: 14/1; 235/1; 290/1; 620/3

Query Match          51.0%  Score 104;  DB 2;  Length 678;

```

Best Local Similarity 50.0%; Pred. NO. 6e-06;
Matches 20; Conservative 10; Mismatches 10; Indels 0; Gaps 0;

QY 2 REGLLAIFKSGGIPALVKMLGSPVDSVLVFYAITTLHNL 41
| | | | | : | | : | | : | | | | |
Db 153 RGGPLLIIFRSGLAEIIRLYDSLVSWHYAVTTLRNLLM 192

RESULT 12
A86157
probable importin alpha subunit [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: A86157
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizlar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C. A.; Li, J. H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: A86157
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-538 <STO>
A:Cross-references: GB:AE005172; NID:g9972381; PIDN:RAG10631.1; GSPDB:CN00141
C:Genetics:
A:Map position: 1
C:Superfamily: pendulin

Query Match 31.9%; Score 65; DB 2; Length 538;
Best Local Similarity 36.8%; Pred. No. 0.86;
Matches 14; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 3 EGLLAIFKSGGIPALVKMLGSPVDVSLFYAITLHLL 40
| : : : | | | : | | | : | : :
Db 279 EKIQTVIDAGVPLRVOLLAHPSFVLIPALRTIGNIV 316
| : : : | | | : | | | : | : :

RESULT 13
T01516
SRP1 protein homolog T10M13.16 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 29-Sep-1999
C:Accession: T01516
R:Johnson, A.F.; de la Bastide, M.; Lodhi, M.; Hoffman, J.; Hasegawa, A.; Gnoj, L.; Gott
Martensen, R.; McCombie, W.
submitted to the EMBL Data Library, May 1997
A:Description: The sequence of the Arabidopsis thaliana T10M13 BAC.
A:Reference number: Z14346
A:Accession: T01516
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-531 <JOH>
A:Cross-references: EMBL:AF001308; NID:g2104523; PID:g2104538
A:Experimental source: cultivar Columbia
C:Genetics:
A:Map position: 4S
A:Introns: 72/3; 106/1; 136/3; 180/3; 242/3; 297/2; 323/3; 356/3; 409/2
A:Note: T10M13.16
C:Superfamily: pendulin

Query Match 31.4%; Score 64; DB 2; Length 531;
Best Local Similarity 35.3%; Pred. No. 1.2;
Matches 12: Conservative 11: Mismatches 11: Indels 0: Gaps 0:

OV 7 AIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 40

[illegible]

RESULT 015
T45588
arm repeat containing protein homolog - Arabidopsis thaliana
N:Alternate names: Protein F12A12.30
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C:Accession: T45588
R:Choišne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.;
submitted to the Protein Sequence Database, December 1999
A:Reference number: 223008

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-660 <CHO>
A:Cross-references: EMBL:AL133314
A:Experimental source: cultivar Columbia; BAC clone F12A12
C:Genetics:
A:Map position: 3
A:Introns: 90/3; 129/3; 279/3
A:Note: "F12A12.30
Query Match 29.9%; Score 61; DB 2; Length 660;
Best Local Similarity 38.2%; Pred. No. 3.7;
Matches: 13; Conservative 9; Mismatches 12; Indels

QY : 6 LAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 39

Dh 388 VATAFAGAIPLVGLLSTPDSRIQHSVTTALLNL 421

Search completed: July 29, 2002, 16:09:10
Job time: 720 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 16:12:06 ; Search time 124.4 seconds
(without alignments)
57.016 Million cell updates/sec

Title: US-09-641-104A-9
Perfect score: 217
Sequence: 1 CSSNKAIVEAGGMOALGLH.....TDPQRVLQVCLTLRLNSD 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	217	100.0	117	6	P79321
2	217	100.0	780	13	Q90424
3	217	100.0	781	13	Q42486
4	211	97.2	781	11	Q9D335
5	195	89.9	769	5	Q9NL44
6	185	85.3	745	4	Q15151
7	185	85.3	745	4	Q9BWC4
8	185	85.3	745	11	P70565
9	185	85.3	773	5	O76152
10	181	83.4	729	13	Q9PVF7
11	173	79.7	806	5	Q25100
12	158	72.8	821	5	O61229
13	77	35.5	811	5	Q18825
14	72.5	33.4	733	4	Q9UBZ1
15	72.5	33.4	1246	4	Q9Y632
16	72.5	33.4	2303	4	O95996

17 70.5 32.5 678 5 O44326
18 67.5 31.1 2274 11 Q9ZLK7
19 65.5 30.2 353 11 Q9K59
20 59.5 27.4 220 4 Q9NH3
21 59.5 27.4 441 4 Q9UDW6
22 59.5 27.4 446 4 Q9UDW5
23 59.5 27.4 499 4 Q9BTW3
24 59.5 27.4 507 4 Q9GT31
25 59.5 27.4 546 11 Q9D3K3
26 59.5 27.4 548 4 Q9GT33
27 59.5 27.4 549 4 Q9GT32
28 59.5 27.4 586 11 Q63648
29 59.5 27.4 590 4 Q9G9Q3
30 59.5 27.4 595 4 Q9G9N0
31 57.5 26.5 539 5 Q9VCL5
32 57 26.3 767 2 Q66008
33 57 26.3 768 2 P94931
34 56.5 26.0 698 10 Q9FNI7
35 56 25.8 158 5 Q9N8C1
36 56 25.8 523 13 Q9PW25
37 56 25.8 534 10 Q9SLX0
38 56 25.8 534 10 Q94GV3
39 56 25.8 660 10 Q9SNC6
40 56 25.8 779 5 Q9W5T9
41 56 25.8 781 5 Q9NHP1
42 55 25.3 198 4 Q9H6L4
43 55 25.3 259 4 Q9UNG9
44 54 24.9 353 5 Q9VR04
45 54 24.9 509 4 Q96AG6

ALIGNMENTS

RESULT 1

P79321 PRELIMINARY: PRT; 117 AA.
AC P79321;
DT 01-MAY-1997 (TRENBLrel. 03, Created)
DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE BETA CATENIN (FRAGMENT).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SMALL INTESTINE;
RA Wintetoe A.K., Fredholm M.;
RT "Evaluation and characterization of a porcine small intestine cDNA library."
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z84131; CAB06327.1; -
DR HSSP; P35222; IG3J.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 2.
DR PROSITE; PS50176; ARM_REPEAT; 2.
FT NON_TER 1 117
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12717 MW; 73F7CCF3B917A41A CRC64;

Query Match 100.0%; Score 217; DB 6; Length 117;
Best Local Similarity 100.0%; Pred. No. 3e-23;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CSSNKAIVEAGGMOALGLHLDTPQRVLQVCLTLRLNSD 41

Db 76 CSSNKAIVEAGGMOALGLHLDTPQRVLQVCLTLRLNSD 116

RESULT 2

Q90424 ID Q90424 PRELIMINARY; PRT; 780 AA.
 AC Q90424;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE B-CATENIN.
 GN CTNNB.
 OS Brachydanio rerio (Zebrafish) (Zebra danio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 OC Cypriniformes; Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96122902; PubMed=8562427;
 RA Kelly G.M., Erezylmaz D.F., Moon R.T.;
 RT "Induction of a secondary embryonic axis in zebrafish occurs following
 the overexpression of beta-catenin.";
 RL Mech. Dev. 53:261-273(1995).
 DR EMBL; U41081; AAC59732.1; -.
 DR HSP; P35222; IG3J.
 DR ZFIN; ZDB-GENE-980526-362; ctnnb.
 DR InterPro; IPR000225; Armadillo.
 DR Pfam; PF00514; Armadillo_seg; 12.
 DR SMART; SM00185; ARM; 11.
 DR PROSITE; PS50176; ARM_REPEAT; 8.
 SQ SEQUENCE 780 AA; 85546 MW; D7A1FB80F94066DC CRC64;

Query Match 100.0%; Score 217; DB 13; Length 780;
 Best Local Similarity 100.0%; Pred. No. 2.6e-22;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CSSNKPAIVEAGGQALGLHLTPSQRVLQVNCNLTWLRNLS 41
 |||||
 Db 349 CSSNKPAIVEAGGQALGLHLTPSQRVLQVNCNLTWLRNLS 389

RESULT 3
 ID 042486 PRELIMINARY; PRT; 781 AA.
 AC 042486;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 13, Last annotation update)
 DE BETA CATENIN.
 GN CHECAT.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN-WHITE LEGRON; TISSUE=DORSAL SKIN;
 RX MEDLINE=97464068; PubMed=9322759;
 RA Lu J., Chung C.M., Wideltz R.B.;
 RT "Isolation and characterization of chicken beta-catenin.";
 RL Gene 196:201-207(1997).
 DR EMBL; U82964; AAB80856.1; -.
 DR HSP; P35222; IG3J.
 DR InterPro; IPR000225; Armadillo.
 DR Pfam; PF00514; Armadillo_seg; 12.
 DR SMART; SM00185; ARM; 11.
 DR PROSITE; PS50176; ARM_REPEAT; 8.
 SQ SEQUENCE 781 AA; 85438 MW; 6D205D9A4DBAC5G2 CRC64;

Query Match 100.0%; Score 217; DB 13; Length 781;
 Best Local Similarity 100.0%; Pred. No. 2.6e-22;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CSSNKPAIVEAGGQALGLHLTPSQRVLQVNCNLTWLRNLS 41
 |||||
 Db 350 CSSNKPAIVEAGGQALGLHLTPSQRVLQVNCNLTWLRNLS 390

RESULT 4
 ID 09D335 PRELIMINARY; PRT; 781 AA.
 AC 09D335;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE ADULT MALE COLON CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY.
 DE CLONE:9030417H18, FULL INSERT SEQUENCE.
 GN CATNB.

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=COLON;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Seito T., Okazaki Y., Gofjbori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schram L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.,
 RA "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 DR EMBL; AK018515; BAB31250.1; -.
 DR HSP; P35222; IG3J.
 DR MGD; MGI:88276; Catnb.
 DR InterPro; IPR000225; Armadillo.
 DR Pfam; PF00514; Armadillo_seg; 12.
 DR SMART; SM00185; ARM; 12.
 DR PROSITE; PS50176; ARM_REPEAT; 7.
 SQ SEQUENCE 781 AA; 85546 MW; 937538C3B5CD75D1 CRC64;

Query Match 97.2%; Score 211; DB 11; Length 781;
 Best Local Similarity 97.6%; Pred. No. 1.9e-21;
 Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CSSNKPAIVEAGGQALGLHLTPSQRVLQVNCNLTWLRNLS 41
 |||||
 Db 350 CSSNKPAIVEAGGQALGLHLTPSQRVLQVNCNLTWLRNLS 390

RESULT 5
 ID 09NL44 PRELIMINARY; PRT; 769 AA.
 AC 09NL44;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE BETA-CATENIN.
 GN CIBETA-CATENIN.
 OS Clona-intestinalis.
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;

```

DR SMART; SM00185; ARM; 9.
DR PROSITE; PS50176; ARM_REPEAT; 7.
SQ SEQUENCE 745 AA; 81744 MW; 3519A0973748BCF4 CRC64;

Query Match 85.3%; Score 185; DB 4; Length 745;
Best Local Similarity 87.8%; Pred. No. 9.1e-18;
Matches 36; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CSSNKPAlVAAGGQALGLHLTDPQSRLVQNCIWTLRNLS 41
DB 341 CPSNKPAlVAAGGQALGLKHLTSNPSRLVQNCIWTLRNLS 381

RESULT 7
Q9BWC4 PRELIMINARY; PRT; 745 AA.
ID Q9BWC4
AC Q9BWC4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE JUNCTION PLAKOGLOBIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI_TaxID=9606;
RX [1]
RN RP SEQUENCE FROM N.A.
RC TISSUE=LUNG; CARCINOMA;
RA Strausberg R.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
[2]
RN RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA; CHORIOCARCINOMA;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC000441; AAH00441.1; -
DR EMBL; BC011865; AAH11865.1; -
DR HSSP; Q02248; 2BCT
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 11.
DR SMART; SM00185; ARM; 12.
DR PROSITE; PS50176; ARM_REPEAT; 7.
SQ SEQUENCE 745 AA; 81726 MW; 34DF7BF4748BCF4 CRC64;

Query Match 85.3%; Score 185; DB 4; Length 745;
Best Local Similarity 87.8%; Pred. No. 9.1e-18;
Matches 36; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CSSNKPAlVAAGGQALGLHLTDPQSRLVQNCIWTLRNLS 41
DB 341 CPSNKPAlVAAGGQALGLKHLTSNPSRLVQNCIWTLRNLS 381

RESULT 8
P70565 PRELIMINARY; PRT; 745 AA.
ID P70565
AC P70565;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PLAKOGLOBIN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
ON NCBI_TaxID=10116;
RX [1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=VAGINA;
RA UneKita Y., Liao S.;
RL "Molecular cloning and sequencing of the rat plakoglobin CDNA.";
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.

```

[2]
RN SEQUENCE FROM N.A.
RP STRAIN-SPRAGUE-DAWLEY; TISSUE-VAGINA;
RA Hiipakka R.A.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U5858; BAB06317.1; -
DR HSP; Q02248; 2BCT.
DR InterPro: IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 11.
DR SMART; SM00185; ARM; 8.
DR PROSITE; PS00176; ARM_REPEAT; 7.
SQ SEQUENCE 745 AA; 81777 MW; 9E2F52910A7ACD41 CRC64;

Query Match 85.3%; Score 185; DB 11; Length 745;
Best Local Similarity 87.8%; Pred. No. 9.1e-18;
Matches 36; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CSSNKPATVEAGGMOALGLHLTDPQSRLVQNCNLTWLRNLS 41
DB 341 CPSNKPATVEAGGMOALGLHLTSPRLVQNCNLTWLRNLS 381

RESULT 9
ID 076152 PRELIMINARY; PRT; 773 AA.
AC 076152;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE BETA-CATENIN.
OS Ciona savignyi.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;
CC Echinoidea; Ciona.
OX NCBI_TaxID=51511;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=98443204; PubMed=9769178;
RA Yoshida S., Marikawa Y., Sato N.;
RT "Regulation of the trunk-tail patterning in the ascidian embryo: a
RT possible interaction of cascades between lithium/beta-catenin and
RT localized maternal factor pem."
RL Dev. Biol. 202;264-279(1998).
DR EMBL; AB012160; BAA32789.1; -
DR HSP; P35222; 1G3J.
DR InterPro: IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 11.
DR SMART; SM00185; ARM; 9.
DR PROSITE; PS00176; ARM_REPEAT; 7.
SQ SEQUENCE 773 AA; 85217 MW; C1340CF82AFEDAB CRC64;

Query Match 85.3%; Score 185; DB 5; Length 773;
Best Local Similarity 85.4%; Pred. No. 9.5e-18;
Matches 35; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 CSSNKPATVEAGGMOALGLHLTDPQSRLVQNCNLTWLRNLS 41
DB 339 CSSNKPATVEAGGMOALGLHLGSRQRLQNCNLTWLRNLS 379

RESULT 10
ID 09PVF7 PRELIMINARY; PRT; 729 AA.
AC 09PVF7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CELL-ADHESION PROTEIN PLAKOGLOBIN.
GN JUP.
OS Brachydanio rerio (zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; teleostei; Euteleostei; Ostariophysi;

OC Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN SEQUENCE FROM N.A.
RP MEDLINE=99386700; PubMed=10456847;
RA Cerda J., Reidenbach S., Pratzel S., Franke W.W.;
RT "cadherin-catenin complexes during zebrafish oogenesis: heterotypic
RT junctions between oocytes and follicle cells."
DR Biol. Reprod. 61;692-704(1999).
DR EMBL; AF099738; AAB56592.1; -
DR HSP; Q02248; 1DOW;
DR ZFIN; ZDB-GENE:991207.22; Jup.
DR InterPro: IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 11.
DR SMART; SM00185; ARM; 8.
DR PROSITE; PS00176; ARM_REPEAT; 8.
SQ SEQUENCE 729 AA; 80033 MW; 91E00417B4FD8CEE CRC64;

Query Match 83.4%; Score 181; DB 13; Length 729;
Best Local Similarity 82.9%; Pred. No. 3.3e-17;
Matches 34; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CSSNKPATVEAGGMOALGLHLTDPQSRLVQNCNLTWLRNLS 41
DB 330 CPSNKPATVEAGGMOALGLHLGSSQRLMONCLWLRNLS 370

RESULT 11
ID 025100 PRELIMINARY; PRT; 806 AA.
AC 025100;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE BETA-CATENIN.
OS Hydra magnipapillata (Hydra).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Anthomedusae;
OC Hydrozoa; Hydra.
OX NCBI_TaxID=6085;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=WILDTYPE 105;
RX MEDLINE=96257271; PubMed=8654977;
RA Hobmayer E., Hatta M., Fischer R., Fujisawa T., Holstein T.W.,
RA Sugiyama T.;
RT "Identification of a Hydra homologue of the beta-
RT catenin/plakoglobin/armadillo gene family."
RL Gene 172;155-159(1996).
DR EMBL; U36781; AAC47137.1; -
DR HSP; Q02248; 2BCT.
DR InterPro: IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 10.
DR SMART; SM00185; ARM; 10.
DR PROSITE; PS00176; ARM_REPEAT; 7.
SQ SEQUENCE 806 AA; 90462 MW; 689E5E982CD5051A CRC64;

Query Match 79.7%; Score 173; DB 5; Length 806;
Best Local Similarity 82.9%; Pred. No. 5.2e-16;
Matches 34; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 CSSNKPATVEAGGMOALGLHLTDPQSRLVQNCNLTWLRNLS 41
DB 409 CSSNKPATVEAGGMOALAHYLSHOSRLVQNCNLTWLRNLS 449

RESULT 12
ID 061229 PRELIMINARY; PRT; 821 AA.
AC 061229;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)

```

RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: U46673; AAC48154.1;
DR EMBL: AF063646; AAC17424.1;
DR HSSP: Q02248; 3BC7.
DR TRANSFAC: T03880;
DR SEQUENCE 811 AA; 92227 MW; 97D6FFDE71BDFFFC CRC64;

Query Match
Best Local Similarity 35.5%; Score 77; DB 5; Length 811;
Matches 19; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 6 PAIVAGGMQALGHLTDPQSRLVQNCILWTLRLNSD 41
Db 335 PNLVAFGGROILANLLSHGSPRLVQSTLETLRNISD 370

RESULT 14
Q9UBZ1 PRELIMINARY; PRT; 733 AA.
ID Q9UBZ1
AC Q9UBZ1
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE APC2 PROTEIN (FRAGMENT).
OS APC2.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RS SEQUENCE FROM N.A.
RP TISSUE=KIDNEY;
RX MEDLINE=99147086; PubMed=10021369;
RA van Es J.H., Kirkpatrick C., van de Wetering M., Molenaar M.,
RT Mies A., Kuipers J., Destree O., Peifer M., Clevers H.;
RT "Identification of APC2, a homologue of the adenomatous polyposis coli
RT tumour suppressor."
RT Curr. Biol. 9:105-108(1999).
RN [2]
RS SEQUENCE FROM N.A.
RP van Es J.H., Kirkpatrick C., van de Wetering M., Molenaar M.,
RT Mies A., Kuipers J., Destree O., Peifer M., Clevers H.;
RT "Adenomatous Polyposis Coli Homologs in Mammals and Flies."
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ012652; CAB61207.1;
DR EMBL: AF128222; AAF01784.1;
DR Interpro: IPR000225; Armadillo.
DR Pfam: PF00514; Armadillo_seg. 7.
DR SMART: SM00185; ARM; 5.
DR NON_TER 733
FT SEQUENCE 733 AA; 80876 MW; 09E56BE5F7032BAD CRC64;

Query Match
Best Local Similarity 33.4%; Score 72.5; DB 4; Length 733;
Matches 21; Conservative 4; Mismatches 15; Indels 25; Gaps 1;

QY 1 CSSNKPATVIRAGG-----MOALGHLTDPQSRLVQNCILWTL 35
Db 588 CQSNLSAIEGSGGILRNVSLSVATRDYRQVLRDHNCLQTLLQHLTSHSLITIVSNACGT 647

QY 36 LRNLIS 40
Db 648 LRNLIS 652

RESULT 15
Q9Y632 PRELIMINARY; PRT; 1246 AA.
ID Q9Y632
AC Q9Y632
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

```

```

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE APC2 PROTEIN (FRAGMENT).
GN APC2.
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Carr I.M., Markham A.F., Colleta P.L., Wai L., Askham J., Morrison E.,
RA Meredith D.M.;
RT "APC2 alternatively spliced cDNA sequence.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -1- COFACTOR: BINDS A 2FE-2S CLUSTER (BY SIMILARITY).
DR EMBL; AF110334; AAD28183.1; -.
DR HSSP; O02248; 3BCT.
DR InterPro; IPR000564; 2Fe2S_ferredoxin.
DR InterPro; IPR000225; Armadillo.
DR InterPro; IPR001818; Matrxin.
DR Pfam; PF00514; Armadillo_seg; 5.
DR PRINTS; PR00159; 2FE2SFRDOXIN.
DR SMART; SM00185; ARM; 3.
DR PROSITE; PS00546; CYSTEINE_SWITCH; UNKNOWN_1.
KW Iron-sulfur.
FT NON_TER 1246 1246
SQ SEQUENCE 1246 AA; 132115 MW; 8C527DA2B3B0BC17 CRC64;

Query Match      33.4%; Score 72.5; DB 4; Length 1246;
Best Local Similarity 32.3%; Pred. No. 0.19;
Matches 21; Conservative 4; Mismatches 15; Indels 25; Gaps 1;

QY 1 CSSNKPAIVEAGG-----MOALGLHLTDPQRLVQNCNLT 35
   | | | | | | | | | | | | | | | | | | | | |
Db 314 CQSNLAIIESGGILRNVSILVATREDYRQVLRDHNCLQTLQHLTSHSLTIVSNACGT 373

QY 36 LRNLS 40
   | | | |
Db 374 LWNLS 378

```

Search completed: July 29, 2002, 16:12:07
 Job time: 477 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 15:57:10 ; Search time 158.47 Seconds
(without alignments)
28.737 Million cell updates/sec

Title: US-09-641-104A-6
Perfect score: 204
Sequence: 1 HREGLAIFKSGIPALVKM.....GSPVDSVLFYAITLHLL 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*
1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Score	Query Match	Length DB ID	Description	
1	204	100.0	41	20	AAV33217 Human beta-catenin
2	204	100.0	781	21	AA07290 Human beta catenin
3	204	100.0	781	21	AAV70740 Human beta-catenin
4	204	100.0	781	22	AA028118 Novel human secret
5	204	100.0	781	22	AAE06038 Mouse beta-catenin
6	204	100.0	781	22	AAE06039 Novel human secret
7	204	100.0	800	22	AAU28306 Novel human secret
8	202	99.0	41	20	AAV33232 Human beta-catenin
9	188	92.2	840	22	AB06196 Drosophila melanog
10	188	92.2	840	22	AB065819 Drosophila melanog
11	188	92.2	840	22	AB065821 Drosophila melanog

12	64	31.4	452	21	AA016853 Arabidopsis thalia
13	64	31.4	452	21	AA045495 Arabidopsis thalia
14	64	31.4	478	21	AA016852 Arabidopsis thalia
15	64	31.4	478	21	AA045494 Arabidopsis thalia
16	64	31.4	531	21	AA016851 Arabidopsis thalia
17	64	31.4	531	21	AA045493 Arabidopsis thalia
18	63	30.9	180	21	AA010593 Arabidopsis thalia
19	63	30.9	180	21	AA018263 Arabidopsis thalia
20	63	30.9	180	21	AA048537 Human beta-catenin
21	61	29.9	41	20	AAV33218 Drosophila melanog
22	61	29.9	442	22	AB065086 Drosophila melanog
23	61	29.9	442	22	AB065412 Zea mays protein f
24	57	27.9	454	21	AA034331 Arabidopsis thalia
25	57	27.9	485	21	AA031001 Arabidopsis thalia
26	57	27.9	486	21	AA031000 Arabidopsis thalia
27	57	27.9	495	21	AA034330 Zea mays protein f
28	57	27.9	527	21	AA034329 Zea mays protein f
29	57	27.9	538	21	AA030999 Arabidopsis thalia
30	57	27.9	779	22	AB066318 Drosophila melanog
31	56.5	27.7	664	22	AA081626 S. epidermidis ope
32	56	27.5	107	22	AA094974 Human reproductive
33	56	27.5	107	22	AAU16365 Human novel secret
34	56	27.5	688	22	AA064567 Human vacuole prot
35	56	27.5	753	22	AAU15912 Human novel peptide
36	55	27.0	116	22	AAU29741 Human polypeptide
37	55	27.0	224	22	AAU29741 Novel human secret
38	55	27.0	543	22	AB071661 Drosophila melanog
39	55	27.0	555	22	AA031943 Amino acid sequenc
40	55	27.0	682	22	AA093210 Human protein sequ
41	54.5	26.7	1110	22	AA036648 Human extracellular
42	54.5	26.7	1192	20	AA023899 Human resenilin bi
43	54.5	26.7	1211	18	AAW24560 Presenilin-interac
44	54.5	26.7	1211	21	AA092336 Human p0071. Homo
45	54	26.5	67	22	AA065659 Human brain expres

ALIGNMENTS

RESULT 1
AAV33217
ID AAV33217 standard; peptide; 41 AA.
AC AAV33217;
DT 18-NOV-1999 (first entry)
XX Human beta-catenin protein armadillo repeat arm3 fragment.

XX Beta-catenin; human; armadillo repeat; treatment; human disease; LEF-1;
XX modulator; interaction domain; transcription factor; TCF-4; oncogenic;
XX tumor suppressor gene; conductin; E-cadherin; screening; carcinoma; APC;
XX Wnt signalling pathway; tumor development; anti-oncogenic; melanoma;
XX organ regeneration; tissue regeneration; hair growth.

OS Homo sapiens.
XX DE19909251-A1.
XX PN
XX XX
XX PD 26-AUG-1999.
XX PF 22-FEB-1999; 99DE-1009251.
XX PR 21-FEB-1998; 98DE-1007390.
XX PA (DELB-) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.
XX PI Birchmeier W, Von Kries J;
XX DR WPI; 1999-470389/40.

XX Agents for treating human diseases, particularly cancer, modulate
PT interaction of beta-catenin with transcription factors or tumor

PT suppressor gene products -
 XX
 PS Disclosure; Page 7; 16pp; German.
 XX
 CC This invention describes a novel agent (A) for treating human disease
 CC which is based on substances (I) that modulate (inhibit or promote) the
 CC interaction of beta-catenin with transcription factors or products of
 CC tumor suppressor genes. The invention also describes (a) peptides (II)
 CC comprising part of the LEF-1/TCF-4 transcription factors, or their
 CC variants and mutants; (b) peptides and related molecules (III) from the
 CC armadillo domain (arm units 3-8) of beta-catenin, and mutants of the
 CC complete beta-catenin molecule, that include at least one of the
 CC specific interaction domains for LEF-1, TCF-4, APC, conductin or
 CC E-cadherin; and (c) enzyme-linked immunosorbent assay (ELISA) for
 CC screening substance libraries for compounds that modulate interaction of
 CC beta-catenin with LEF-1/TCF, APC, conductin or E-cadherin. Beta-catenin
 CC is a key compound in the Wnt signalling pathway and is involved in
 CC development of tumors. Generally its interaction with LEF-1 or TCF-4 is
 CC oncogenic but interaction with APC, conductin or E-cadherin is
 CC anti-oncogenic. (A) which inhibit interaction are particularly used to
 CC treat tumors, especially carcinoma of the colon and melanoma, but also,
 CC where they promote interaction, to stimulate regeneration of organs and
 CC tissues, specifically hair growth. AAY33217-Y33222 represent human
 CC beta-catenin armadillo repeat fragments described in the method of the
 CC invention.
 XX
 SQ Sequence 41 AA;

Query Match 100.0%; Score 204; DB 20; Length 41;
 Best Local Similarity 100.0%; Pred. No. 4.2e-23;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHLL 41
 DB 1 hregllaifksggipalvkmlgspvdsvlfyaittlhlll 41

RESULT 2
 AAB07290
 ID AAB07290 standard; Protein; 781 AA.
 XX
 AC AAB07290;
 XX
 DT 01-NOV-2000 (first entry)
 XX
 DE Human beta catenin.
 XX
 KW Beta catenin; cadherin; metastasis; cadherin-associated protein;
 KW human; colorectal cancer; melanoma; antisense oligonucleotide;
 KW gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN US6066500-A.
 XX
 PD 23-MAY-2000.
 XX
 PF 25-JUN-1999; 99US-0344519.
 XX
 PR 25-JUN-1999; 99US-0344519.
 XX
 PS (ISIS-) ISIS PHARM INC.
 PA Bennett CF, Cowser LM;
 PI WPI; 2000-410651/35.
 XX N-PSDB; AAS58320.
 DR
 XX New antisense compounds targeting nucleic acids encoding human beta
 PT catenin (HBC) useful for treating diseases associated with HBC
 PT expression and as prophylaxis to prevent or delay infection,
 PT inflammation or tumor formation

XX
 PS Example 13; Columns 45-52; 35pp; English.
 XX
 CC Beta catenin is a member of the catenin family of cytosolic proteins and
 CC a key member of the Wnt signalling pathway. Catenins interact with the
 CC cytoplasmic domains of cadherin glycoproteins, and are important in
 CC maintaining cell adhesiveness. The loss of cell adhesiveness is
 CC implicated in metastasis. Beta catenin is also known as
 CC cadherin-associated protein and is implicated in colorectal cancer and
 CC melanoma. The present sequence is the human beta catenin protein. The
 CC coding sequence of this protein was used in the present invention to
 CC design antisense oligonucleotides (AAS58327-A58366). The
 CC oligonucleotides are capable of hybridizing to human beta catenin, in
 CC order to inhibit expression of human beta catenin. The oligonucleotides
 CC may be used in gene therapy for colorectal cancer or melanoma.
 XX
 SQ Sequence 781 AA;

Query Match 100.0%; Score 204; DB 21; Length 781;
 Best Local Similarity 100.0%; Pred. No. 1.7e-21;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHLL 41
 DB 224 hregllaifksggipalvkmlgspvdsvlfyaittlhlll 264

RESULT 3
 AAY70740
 ID AAY70740 standard; protein; 781 AA.
 XX
 AC AAY70740;
 XX
 DT 24-JUL-2000 (first entry)
 XX
 DE Human beta-catenin.
 XX
 KW Human beta-catenin; Wnt antagonist; contraceptive; contraceptive vaccine;
 KW oocyte development; female primate contraception; oocyte viability;
 KW monoclonal antibody; Wnt signalling.
 XX
 OS Homo sapiens.
 XX
 PN WO200021555-A1.
 XX
 PD 20-APR-2000.
 XX
 PF 13-OCT-1999; 99WO-US23640.
 XX
 PR 15-OCT-1998; 98US-0104355.
 XX
 PS (HARD) HARVARD COLLEGE.
 PA McMahon AP, Parr BA, Vaino S;
 PI WPI; 2000-317845/27.
 XX
 DR Contraceptive composition for inhibiting oocyte development in a female
 XX primate comprises a Wnt polypeptide antagonist
 PT
 PS Example 3; Page 26; 57pp; English.
 XX
 CC The patent discloses a method of female primate contraception comprising
 CC administering an antagonist of a Wnt polypeptide, inhibiting oocyte
 CC development. Wnt polypeptides are useful for promotive maturation of an
 CC immature oocyte. Wnt polypeptides are also useful for increasing the
 CC number of mature oocytes and to enhance oocyte viability. Soluble
 CC fragments of Wnt polypeptides have the ability to inhibit Wnt signalling,
 CC e.g., by blocking binding of a naturally-occurring Wnt protein to its
 CC receptor. They may be used to generate monoclonal antibodies which can
 CC inhibit oocyte development. The present sequence is the human
 CC beta-catenin protein. Dominant negative mutants of beta-catenin lack one

CC or more armidillo-like repeats which participate in cadherin binding.
 CC Other mutants include those lacking amino acids 555-781 or 424-781,
 CC or 1-422 of beta-catenin.
 XX
 SQ Sequence 781 AA;

Query Match 100.0%; Score 204; DB 21; Length 781;
 Best Local Similarity 100.0%; Pred. No. 1.7e-21;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 41
 DB 224 hregllalifksggipalvkmigspvdsvlfyaittlhnl 264

RESULT 4
 AAU28118
 ID AAU28118 standard; Protein: 781 AA.
 XX
 AC AAU28118;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Novel human secretory protein, Seq ID No 287.
 XX
 KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
 KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
 KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
 KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
 KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
 KW gut protection; lung; liver fibrosis; immune deficiency; infection;
 KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
 KW fertility; analgesic; pain; antigen.
 XX
 OS Homo sapiens.
 XX
 PN WO200166689-A2.
 XX
 PD 13-SEP-2001.
 XX
 PF 05-MAR-2001; 2001WO-US04942.
 XX
 PR 07-MAR-2000; 2000US-0519705.
 PR 19-MAY-2000; 2000US-0574454.
 PR 17-JUN-2000; 2000US-0596193.
 PR 14-JUL-2000; 2000US-0616847.
 PR 19-SEP-2000; 2000US-0665363.
 PR 20-OCT-2000; 2000US-0693267.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
 PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
 XX
 DR WPI; 2001-589934/66.
 DR N-PSDB; AAS45018.
 XX
 PT Novel polypeptides and nucleic acids obtained from cDNA libraries
 PT prepared from various human tissues, for diagnosis and treatment of
 PT cancer, neurological, inflammatory, and autoimmune disorders -
 XX
 PS Example 4; SEQ ID No 287; 107pp; English.
 XX
 CC The invention relates to novel isolated human secreted polypeptides (I)
 CC and polynucleotides (II). (I) and (II) are useful for treating
 CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
 CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
 CC involved in increasing haematopoiesis, stem cell survival, bone growth
 CC and remodeling. (I), (II) and modulators of (II) are useful for
 CC prophylaxis or treatment of one or more cancers. (II) is also useful for
 CC creating transgenic animals useful for studying the in vivo activities of

CC the polypeptide as well as for studying modulators of the polypeptides.
 CC (I) induces the proliferation of neural cells and regeneration of nerve
 CC and brain tissue and is useful for the treatment of central and
 CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
 CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
 CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
 CC activity, regulation of haematopoiesis and is useful for treating myeloid
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 CC tissue growth, and in tissue repair, healing of burns, incisions,
 CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
 CC disorders, or periodontal disease. Furthermore, (I) is also useful for
 CC gut protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues, various immune deficiencies and
 CC disorders including severe combined immunodeficiency (SCID), bacterial or
 CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
 CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
 CC reactions and conditions, such as asthma or other respiratory problems.
 CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
 CC fertility, metabolism, catabolism, anabolism, storage or elimination of
 CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
 CC analgesic effects or other pain reducing effects, immunoglobulin like
 CC activity and can act as an antigen in a vaccine composition to raise an
 CC immune response. AAU28020-AAU28395 represent novel human secreted protein
 CC amino acid sequences of the invention.
 XX
 SQ Sequence 781 AA;

Query Match 100.0%; Score 204; DB 22; Length 781;
 Best Local Similarity 100.0%; Pred. No. 1.7e-21;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 41
 DB 224 hregllalifksggipalvkmigspvdsvlfyaittlhnl 264

RESULT 5
 AAU28038
 ID AAU28038 standard; Protein: 781 AA.
 XX
 AC AAU28038;
 XX
 DT 25-SEP-2001 (first entry)
 XX
 DE Human beta-catenin protein #1.
 XX
 KW Human; stem cell culturing; progenitor cell; pluripotential phenotype;
 KW transplantation; haematopoietic function; allogeneic recipient;
 KW signalling pathway; beta-catenin.
 XX
 OS Homo sapiens.
 XX
 PN WO200152649-A1;
 XX
 PD 26-JUL-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01459.
 XX
 PR 18-JAN-2000; 2000US-0176786.
 XX
 PA (STRD-) UNIV LELAND STANFORD JUNIOR.
 XX
 PI Reya T, Nusse R, Weissman IL;
 XX
 DR WPI; 2001-465328/50.
 DR N-PSDB; AAD11164.
 XX
 PT In-vitro expansion of mammalian stem or progenitor cells, useful for
 PT producing cells that retain their pluripotential phenotype after
 PT expansion for use in transplantation, involves increasing the levels of
 PT beta-catenin in the cell -

Tue Jul 30 08:15:34 2002

us-09-641-104a-6.rag

XX

PS Disclosure; Page 23-24; 33pp; English.

XX

CC The present invention relates to a method for in vitro expansion of
 CC mammalian stem or progenitor cells, which comprising increasing the
 CC intracellular concentration of beta-catenin in a progenitor or stem cell
 CC in an in vitro culture medium for a period sufficient for the progenitor
 CC or stem cell to divide. The number of cells having the functional
 CC phenotype of the stem or progenitor cells is expanded. The method is
 CC useful for culturing stem cells and progenitor cells, which retain their
 CC pluripotential phenotype after expansion, in vitro. The expanded cell
 CC populations are useful as a source of stem cells, e.g. to reconstitute
 CC function in a host that is deficient in a particular cell lineage or
 CC lineages. The expanded cell populations are also useful in
 CC transplantation to restore haematopoietic function to autologous or
 CC allogeneic recipients. The present sequence is human beta-catenin
 CC protein. Beta-catenin is a pivotal player in the signalling pathway
 CC initiated by Wnt proteins, which are mediators of several developmental
 CC processes.

XX

SQ Sequence 781 AA;

Query Match 100.0%; Score 204; DB 22; Length 781;

Best Local Similarity 100.0%; Pred. No. 1.7e-21;

Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHMLL 41

|||||

224 hregllaifksggipalvkmlgspvdsvlfyaittlhml 264

Db

RESULT 6

AAE06039

ID AAE06039 standard; Protein; 781 AA.

XX

AC AAE06039;

XX

DT 25-SEP-2001 (first entry)

XX

DE Mouse beta-catenin protein #2.

XX

KW Mouse; stem cell culturing; progenitor cell; pluripotential phenotype;
 KW transplantation; haematopoietic function; allogeneic recipient;
 KW signalling pathway; beta-catenin.

XX

OS Mus musculus.

XX

XX WO200152649-A1.

XX

XX 26-JUL-2001.

XX

XX 17-JAN-2001; 2001WO-US01459.

XX

XX 18-JAN-2000; 2000US-0176786.

XX

XX (STRD) UNIV LELAND STANFORD JUNIOR.

XX

PA Reya T, Nusse R, Weissman IL;

XX

XX WPI; 2001-465328/50.

XX

XX N-PSDB; AAD11165.

XX

PT In vitro expansion of mammalian stem or progenitor cells, useful for
 PT producing cells that retain their pluripotential phenotype after
 PT expansion for use in transplantation, involves increasing the levels of
 PT beta-catenin in the cell -

XX

PS Disclosure; Page 28-30; 33pp; English.

XX

CC The present invention relates to a method for in vitro expansion of
 CC mammalian stem or progenitor cells, which comprising increasing the
 CC intracellular concentration of beta-catenin in a progenitor or stem cell

XX

XX

XX

XX

XX

XX

CC in an in vitro culture medium for a period sufficient for the progenitor
 CC or stem cell to divide. The number of cells having the functional
 CC phenotype of the stem or progenitor cells is expanded. The method is
 CC useful for culturing stem cells and progenitor cells, which retain their
 CC pluripotential phenotype after expansion, in vitro. The expanded cell
 CC populations are useful as a source of stem cells, e.g. to reconstitute
 CC function in a host that is deficient in a particular cell lineage or
 CC lineages. The expanded cell populations are also useful in
 CC transplantation to restore haematopoietic function to autologous or
 CC allogeneic recipients. The present sequence is mouse beta-catenin
 CC protein. Beta-catenin is a pivotal player in the signalling pathway
 CC initiated by Wnt proteins, which are mediators of several developmental
 CC processes.

SQ Sequence 781 AA;

Query Match 100.0%; Score 204; DB 22; Length 781;

Best Local Similarity 100.0%; Pred. No. 1.7e-21;

Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHMLL 41

|||||

224 hregllaifksggipalvkmlgspvdsvlfyaittlhml 264

Db

RESULT 7

AAU28306

ID AAU28306 standard; Protein; 800 AA.

XX

XX AAU28306;

XX

DT 18-DEC-2001 (first entry)

XX

DE Novel human secretory protein, Seq ID No 663.

XX

KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
 KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
 KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
 KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
 KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
 KW gut protection; lung; liver fibrosis; immune deficiency; infection;
 KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
 KW fertility; analgesic; pain; antigen.

XX

OS Homo sapiens.

XX

XX WO200166689-A2.

XX

XX 13-SEP-2001.

XX

XX 05-MAR-2001; 2001WO-US04942.

XX

XX 07-MAR-2000; 2000US-0519705.

XX

XX 19-MAY-2000; 2000US-0574454.

XX

XX 17-JUN-2000; 2000US-0596193.

XX

XX 14-JUL-2000; 2000US-0616847.

XX

XX 19-SEP-2000; 2000US-0665363.

XX

XX 20-OCT-2000; 2000US-0693267.

XX

XX (HYSEQ) HYSEQ INC.

XX

XX Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;

XX

XX Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;

XX

XX WPI; 2001-589934/66.

XX

XX N-PSDB; AAS45206.

XX

XX Novel polypeptides and nucleic acids obtained from cDNA libraries
 XX prepared from various human tissues, for diagnosis and treatment of
 XX cancer, neurological, inflammatory, and autoimmune disorders -

XX

Example 2; SEQ ID NO 663; 107pp; English.

The invention relates to novel isolated human secreted polypeptides (I) and polynucleotides (II). (I) and (II) are useful for treating inflammatory conditions such as arthritis, nephritis, Crohn's disease, ischaemia-reperfusion injury, shock, sepsis, immune responses, and is involved in increasing haematopoiesis, stem cell survival, bone growth and remodeling. (I), (II) and modulators of (II) are useful for prophylaxis or treatment of one or more cancers. (II) is also useful for creating transgenic animals useful for studying the in vivo activities of the polypeptide as well as for studying modulators of the polypeptides. (I) induces the proliferation of neural cells and regeneration of nerve and brain tissue and is useful for the treatment of central and peripheral nervous system diseases and neuropathies, such as Alzheimer's, Parkinson's disease, Huntington's disease, and amyotrophic lateral sclerosis. In addition, (I) is involved in chemotactic or chemokinetic activity, regulation of haematopoiesis and is useful for treating myeloid or lymphoid cell disorders, platelet disorders such as thrombocytopenia and for regeneration of bone, cartilage, tendon, ligament and/or nerve tissue growth, and in tissue repair, healing of burns, incisions, ulcers, for treating osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease. Furthermore, (I) is also useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, various immune deficiencies and disorders including severe combined immunodeficiency (SCID), bacterial or fungal infections, autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic reactions and conditions, such as asthma or other respiratory problems. In addition, (I) affects biorhythms or circadian cycles of rhythms, fertility, metabolism, catabolism, anabolism, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides analgesic effects or other pain reducing effects, immunoglobulin like activity and can act as an antigen in a vaccine composition to raise an immune response. AAU28020-AAU28395 represent novel human secreted protein amino acid sequences of the invention.

Sequence 800 AA;

Query Match 100.0%; Score 204; DB 22; Length 800;
Best Local Similarity 100.0%; Pred. No. 1.8e-21;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 41
Db 237 hregllaifksggipalvkmlgspvdsvlfyaittlhnl 277

RESULT 8

AAU33232
ID AAY33232 standard; peptide; 41 AA.

AC AAY33232;

DT 18-NOV-1999 (first entry)

DE Human beta-catenin protein mutant armadillo repeat arm 3.

KW Beta-catenin; human; armadillo repeat; treatment; human disease; LEF-1; modulator; interaction domain; transcription factor; TCF-4; oncogenic; tumor suppressor gene; conductin; E-cadherin; screening; carcinoma; APC; Wnt signalling pathway; tumor development; anti-oncogenic; melanoma; organ regeneration; tissue regeneration; hair growth; mutant.

OS Homo sapiens.

OS Synthetic.

PN DE19909251-A1.

PD 26-AUG-1999.

XX 22-FEB-1999; 99DE-1009251.

PR 21-FEB-1998; 98DE-1007390.
XX (DELB;) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.

PI Birschmeier W, Von Kries J;

XX WPI: 1999-470389/40.

PT Agents for treating human diseases, particularly cancer, modulate interaction of beta-catenin with transcription factors or tumor suppressor gene products

XX Example 4; Fig 5; 16pp; German.

XX This invention describes a novel agent (A) for treating human disease which is based on substances (I) that modulate (inhibit or promote) the interaction of beta-catenin with transcription factors or products of tumor-suppressor genes. The invention also describes (a) peptides (II) comprising part of the LEF-1/TCF-4 transcription factors, or their variants and mutants; (b) peptides and related molecules (III) from the armadillo domain (arm units 3-8) of beta-catenin, and mutants of the complete beta-catenin molecule, that include at least one of the specific interaction domains for LEF-1, TCF-4, APC, conductin or E-cadherin; and (c) enzyme-linked immunosorbent assay (ELISA) for screening substance libraries for compounds that modulate interaction of beta-catenin with LEF-1/TCF-4, APC, conductin or E-cadherin. Beta-catenin is a key compound in the Wnt signalling pathway and is involved in development of tumors. Generally its interaction with LEF-1 or TCF-4 is oncogenic but interaction with APC, conductin or E-cadherin is anti-oncogenic. (A) which inhibit interaction are particularly used to treat tumors, especially carcinoma of the colon and melanoma, but also, where they promote interaction, to stimulate regeneration of organs and tissues, specifically hair growth. AAY33230-Y33241 represent mutant human beta-catenin armadillo repeat fragments described in the method of the invention.

XX Sequence 41 AA;

Query Match 99.0%; Score 202; DB 20; Length 41;
Best Local Similarity 97.6%; Pred. No. 8.3e-23;
Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 41
Db 21 hregllaifksggipalvkmlgspvdsvlfyaittlhnl 41

RESULT 9

ABB60196
ID ABB60196 standard; Protein; 840 AA.

AC ABB60196;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 7380.

KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical;

OS Drosophila melanogaster.

PN WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-0509231.

XX 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE-) PE CORP NY.

XX	Venter JC, Adams M, Li PWD, Myers EW;
PI	
XX	WPI; 2001-656860/75.
DR	N-PSDB; ABL04299.
XX	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more
PT	genes from Drosophila and for elucidating cell signalling and cell-cell
PT	interactions -
XX	
XX	Disclosure; SEQ ID NO 7380; 21pp + Sequence Listing; English.
PS	
XX	
XX	The invention relates to an isolated nucleic acid detection reagent
CC	capable of detecting 1000 or more genes from Drosophila. The invention is
CC	useful in developmental biology and in elucidating cell signalling and
CC	cell-cell interactions in higher eukaryotes for the development of
CC	insecticides, therapeutics and pharmaceutical drugs. The invention
CC	discloses genomic DNA sequences (ABL61676-ABL30511), expressed DNA
CC	sequences (ABL01840-ABL16175) and the encoded proteins
CC	(ABB57737-ABB72072).
CC	The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences .
XX	
XX	Sequence 840 AA;
SQ	

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA-sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072).

The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at: ftp.wipo.int/pub/published_pct_sequences.

Sequence 840 AA:

Query Match : 92.2%; Score 188; DB 22; Length 840;
Best Local Similarity 90.2%; Pred. NO. 4.6e-19;
Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY .1 HREGILATFKSGIPALVKMLGSPVDSVLFAITTLHMLLL 41
Db 229 hrqgllaifksggipalvklisspsvesvlfaittlhlml 269

RESULT. 11.
ABB65821
ID ABB65821 standard; Protein; 840 AA.
XX AC ABB65821;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster polypeptide SEQ ID NO 24255.
XX KW Drosophila; developmental biology; cell signalling; Insecticide; pharmaceutical.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US09231.
XX PR 23-MAR-2000; 2000US-191637P.
XX BR 11-JUL-2000; 2000US-061415O.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX DR WPI; 2001-656860/75.
XX DR N-PSDB; ABL09924.
XX PT New, isolated nucleic acid detection reagent for detecting 1000 or more genes, from Drosophila and for elucidating cell signalling and cell-cell interactions
XX PS Disclosure; SEQ ID NO 24255; Zipp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072).
XX CC The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX CC

SO Sequence 840 AA;

Query Match 92.2%; Score 188; DB 22; Length 840;
Best Local Similarity 90.2%; Pred. No. 4,6e-19;
Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HREGLLATFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 41
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 229 hrqgllaifksggipalvklisspveslvyaittlhnl 269

RESULT 12
AAG16853
ID AAG16853 standard; Protein; 452 AA.
XX AC AAG16853;
XX
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 17655.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EPI033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0133180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 29-MAR-1999; 99US-0126264.
XX PR 01-APR-1999; 99US-0126785.
XX PR 06-APR-1999; 99US-0127462.
XX PR 08-APR-1999; 99US-0128234.
XX PR 16-APR-1999; 99US-0128714.
XX PR 19-APR-1999; 99US-0129845.
XX PR 21-APR-1999; 99US-0130077.
XX PR 23-APR-1999; 99US-0130449.
XX PR 28-APR-1999; 99US-0130510.
XX PR 30-APR-1999; 99US-0130891.
XX PR 30-APR-1999; 99US-0131449.
XX PR 04-MAY-1999; 99US-0132048.
XX PR 05-MAY-1999; 99US-0132407.
XX PR 06-MAY-1999; 99US-0132484.
XX PR 06-MAY-1999; 99US-0132485.
XX PR 07-MAY-1999; 99US-0132486.
XX PR 11-MAY-1999; 99US-0132487.
XX PR 14-MAY-1999; 99US-0132863.
XX PR 14-MAY-1999; 99US-0134218.
XX PR 14-MAY-1999; 99US-0134219.
XX PR 14-MAY-1999; 99US-0134221.
XX PR 18-MAY-1999; 99US-0134370.
XX PR 19-MAY-1999; 99US-0134768.
XX PR 20-MAY-1999; 99US-0134941.
XX PR 21-MAY-1999; 99US-0135124.
XX PR 24-MAY-1999; 99US-0135353.
XX PR 25-MAY-1999; 99US-0135629.
XX PR 27-MAY-1999; 99US-0136021.
XX PR 28-MAY-1999; 99US-0136392.
XX PR 01-JUN-1999; 99US-0136782.
XX PR 03-JUN-1999; 99US-0137222.
XX PR 04-JUN-1999; 99US-0137528.
XX PR 06-AUG-1999; 99US-0137502.
XX PR 07-JUN-1999; 99US-0137724.
XX PR 08-JUN-1999; 99US-0138094.
XX PR 10-JUN-1999; 99US-0138540.
XX PR 10-JUN-1999; 99US-0138847.
XX PR 14-JUN-1999; 99US-0139119.
XX PR 16-JUN-1999; 99US-0139452.
XX PR 16-JUN-1999; 99US-0139453.
XX PR 17-JUN-1999; 99US-0139492.
XX PR 18-JUN-1999; 99US-0139454.
XX PR 18-JUN-1999; 99US-0139455.
XX PR 18-JUN-1999; 99US-0139456.
XX PR 18-JUN-1999; 99US-0139457.
XX PR 18-JUN-1999; 99US-0139458.
XX PR 18-JUN-1999; 99US-0139459.
XX PR 18-JUN-1999; 99US-0139460.
XX PR 18-JUN-1999; 99US-0139461.
XX PR 18-JUN-1999; 99US-0139462.
XX PR 18-JUN-1999; 99US-0139463.
XX PR 18-JUN-1999; 99US-0139750.
XX PR 18-JUN-1999; 99US-0139763.
XX PR 21-JUN-1999; 99US-0139817.
XX PR 22-JUN-1999; 99US-0139899.
XX PR 23-JUN-1999; 99US-0140353.
XX PR 24-JUN-1999; 99US-0140354.
XX PR 24-JUN-1999; 99US-0140695.
XX PR 28-JUN-1999; 99US-0140823.
XX PR 29-JUN-1999; 99US-0140991.
XX PR 30-JUN-1999; 99US-0141287.
XX PR 01-JUL-1999; 99US-0141842.
XX PR 01-JUL-1999; 99US-0142154.
XX PR 02-JUL-1999; 99US-0142055.
XX PR 06-JUL-1999; 99US-0142390.
XX PR 08-JUL-1999; 99US-0142803.
XX PR 09-JUL-1999; 99US-0142920.
XX PR 12-JUL-1999; 99US-0142977.
XX PR 13-JUL-1999; 99US-0143542.
XX PR 14-JUL-1999; 99US-0143624.
XX PR 15-JUL-1999; 99US-0144005.
XX PR 16-JUL-1999; 99US-0144085.
XX PR 16-JUL-1999; 99US-0144086.
XX PR 19-JUL-1999; 99US-0144325.
XX PR 19-JUL-1999; 99US-0144331.
XX PR 19-JUL-1999; 99US-0144332.
XX PR 19-JUL-1999; 99US-0144333.
XX PR 19-JUL-1999; 99US-0144334.
XX PR 19-JUL-1999; 99US-0144335.
XX PR 20-JUL-1999; 99US-0144352.
XX PR 20-JUL-1999; 99US-0144632.
XX PR 20-JUL-1999; 99US-0144884.
XX PR 21-JUL-1999; 99US-0144814.
XX PR 21-JUL-1999; 99US-0145086.
XX PR 21-JUL-1999; 99US-0145088.
XX PR 22-JUL-1999; 99US-0145085.
XX PR 22-JUL-1999; 99US-0145087.
XX PR 22-JUL-1999; 99US-0145089.
XX PR 22-JUL-1999; 99US-0145192.
XX PR 23-JUL-1999; 99US-0145145.
XX PR 23-JUL-1999; 99US-0145218.
XX PR 26-JUL-1999; 99US-0145224.
XX PR 27-JUL-1999; 99US-0145276.
XX PR 27-JUL-1999; 99US-0145913.
XX PR 27-JUL-1999; 99US-0145918.
XX PR 28-JUL-1999; 99US-0145919.
XX PR 02-AUG-1999; 99US-0145951.
XX PR 02-AUG-1999; 99US-0146386.
XX PR 02-AUG-1999; 99US-0146388.
XX PR 02-AUG-1999; 99US-0146389.
XX PR 03-AUG-1999; 99US-0147038.
XX PR 04-AUG-1999; 99US-0147204.
XX PR 04-AUG-1999; 99US-0147302.
XX PR 05-AUG-1999; 99US-0147192.
XX PR 05-AUG-1999; 99US-0147260.
XX PR 06-AUG-1999; 99US-0147303.
XX PR 06-AUG-1999; 99US-0147416.

09-AUG-1999; 99US-0147493.
09-AUG-1999; 99US-0147935.
10-AUG-1999; 99US-0148171.
11-AUG-1999; 99US-0148319.
12-AUG-1999; 99US-0148341.
13-AUG-1999; 99US-0148565.
13-AUG-1999; 99US-0148684.
16-AUG-1999; 99US-0149368.
17-AUG-1999; 99US-0149175.
18-AUG-1999; 99US-0149426.
20-AUG-1999; 99US-0149722.
20-AUG-1999; 99US-0149723.
20-AUG-1999; 99US-0149929.
23-AUG-1999; 99US-0149902.
23-AUG-1999; 99US-0149930.
25-AUG-1999; 99US-0150566.
26-AUG-1999; 99US-0150884.
27-AUG-1999; 99US-0151065.
27-AUG-1999; 99US-0151066.
27-AUG-1999; 99US-0151080.
30-AUG-1999; 99US-0151303.
31-AUG-1999; 99US-0151436.
01-SEP-1999; 99US-0151930.
07-SEP-1999; 99US-0152363.
10-SEP-1999; 99US-0153070.
13-SEP-1999; 99US-0153758.
15-SEP-1999; 99US-0154018.
16-SEP-1999; 99US-0154039.
20-SEP-1999; 99US-0154779.
22-SEP-1999; 99US-0155139.
23-SEP-1999; 99US-0155486.
24-SEP-1999; 99US-0155659.
28-SEP-1999; 99US-0156458.
29-SEP-1999; 99US-0156596.
04-OCT-1999; 99US-0157117.
05-OCT-1999; 99US-0157753.
06-OCT-1999; 99US-0157865.
07-OCT-1999; 99US-0158029.
08-OCT-1999; 99US-0158232.
12-OCT-1999; 99US-0158369.
13-OCT-1999; 99US-0159293.
13-OCT-1999; 99US-0159294.
14-OCT-1999; 99US-0159329.
14-OCT-1999; 99US-0159330.
14-OCT-1999; 99US-0159331.
14-OCT-1999; 99US-0159637.
14-OCT-1999; 99US-0159638.
18-OCT-1999; 99US-0159584.
21-OCT-1999; 99US-0160741.
21-OCT-1999; 99US-0160767.
21-OCT-1999; 99US-0160768.
21-OCT-1999; 99US-0160770.
21-OCT-1999; 99US-0160814.
21-OCT-1999; 99US-0160815.
22-OCT-1999; 99US-0160980.
22-OCT-1999; 99US-0160981.
22-OCT-1999; 99US-0160989.
25-OCT-1999; 99US-0161404.
25-OCT-1999; 99US-0161405.
25-OCT-1999; 99US-0161406.
26-OCT-1999; 99US-0161359.
26-OCT-1999; 99US-0161360.
26-OCT-1999; 99US-0161361.
28-OCT-1999; 99US-0161920.
28-OCT-1999; 99US-0161992.
28-OCT-1999; 99US-0161993.
29-OCT-1999; 99US-0162142.
29-OCT-1999; 99US-0162142.

Query Match 31.4%; Score 64; DB 21; Length 452;
Best Local Similarity 35.3%; Pred. No. 0.6;
Matches 12; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

QY 17 AIFKSGIPALVMLGSPVDVLEVAITTLHLL 40
Db 203 avieagvprliqlighsspsvllipaltigniv 236
RESULT 13
AAG45495; 452 AA.
XX AAG45495;
XX AC AAG45495;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 57125.
XX KW protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EPI033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 21-MAY-1999; 99US-0135629.
PR 24-MAY-1999; 99US-0136021.
PR 25-MAY-1999; 99US-0136392.
PR 27-MAY-1999; 99US-0136782.
PR 28-MAY-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.

PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query match 31.4%; Score 64; DB 21; Length 452;
Best Local Similarity 35.3%; Pred. No. 0.6;
Matches 12; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

Qy 7 AIFKSGGIPALVKMLGSPVDSVLFVYAITTLHNL 40
| : : : | : : : | : : : | : : : | : : : | : : :
Db 203 avicagvprliqlghspsvlpaltigniv 236
| : : : | : : : | : : : | : : : | : : : | : : :
RESULT 14
AAG16852.?

PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 22-JUL-1999; 99US-0145194.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.

PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 31.4%; Score 64; DB 21; Length 478;
Best Local Similarity 35.3%; Pred. No. 0.65;
Matches 12; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

QY 7-AIFKSGIPALVKMLGSPVDSVLFVAITTLHNL 40
Db 229-avieagvprliqlighspsvliipalrtigniv 262

Search completed: July 29, 2002, 16:07:52
Job time: 642 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 15:57:10 ; Search time 57.6 seconds
(without alignments)
17.386 Million cell updates/sec

Title: US-09-641-104A-6
Perfect score: 204
Sequence: 1 HREGLAIPKSGIPALVKM.....GSPVDSVLFYAITTLNLL 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	ID	Description
1	53.5	26.2	620	US-08-982-785A-10	Sequence 10, Appl
2	53.5	26.2	682	US-08-982-785A-9	Sequence 9, Appl
3	53	26.0	433	US-08-883-515-2	Sequence 2, Appl
4	52	25.5	44	US-08-982-785A-3	Sequence 3, Appl
5	52	25.5	521	US-08-933-227-1	Sequence 1, Appl
6	52	25.5	521	US-08-933-227-3	Sequence 3, Appl
7	52	25.5	529	US-08-933-227-4	Sequence 4, Appl
8	51.5	25.2	686	US-08-982-785A-8	Sequence 8, Appl
9	51.5	25.2	756	US-08-982-785A-2	Sequence 2, Appl
10	51	25.0	666	US-08-982-785A-11	Sequence 11, Appl
11	51	25.0	856	US-08-486-099-103	Sequence 103, App
12	51	25.0	856	US-08-484-223B-103	Sequence 103, App
13	51	25.0	856	US-08-919-597-103	Sequence 103, App
14	51	25.0	856	US-08-475-668A-103	Sequence 103, App
15	51	25.0	856	US-08-485-551A-103	Sequence 103, App
16	51	25.0	856	US-08-471-913A-103	Sequence 103, App
17	51	25.0	856	US-08-485-264A-103	Sequence 103, App
18	51	25.0	856	US-08-474-349A-103	Sequence 103, App
19	51	25.0	857	US-08-220-151-10	Sequence 10, Appl
20	51	25.0	857	US-08-413-118-10	Sequence 10, Appl
21	51	25.0	857	US-08-804-439A-18	Sequence 18, Appl
22	51	25.0	857	US-08-360-107A-113	Sequence 113, App
23	51	25.0	857	US-08-473-446-10	Sequence 10, Appl
24	51	25.0	857	US-08-720-229-18	Sequence 18, Appl
25	51	25.0	2842	US-07-741-940-7	Sequence 7, Appl
26	51	25.0	2842	US-08-289-548A-7	Sequence 7, Appl
27	51	25.0	2842	US-08-452-654-7	Sequence 7, Appl

28	51	25.0	2843	1	US-07-741-940-2	Sequence 2, Appl
29	51	25.0	2843	1	US-08-289-548A-2	Sequence 2, Appl
30	51	25.0	2843	1	US-08-452-654-2	Sequence 2, Appl
31	51	25.0	2843	1	US-08-452-655B-2	Sequence 2, Appl
32	51	25.0	2843	1	US-08-452-655B-7	Sequence 7, Appl
33	51	25.0	2843	2	US-08-370-235A-2	Sequence 2, Appl
34	51	25.0	2843	3	US-08-450-582-2	Sequence 2, Appl
35	51	25.0	2843	3	US-08-450-582-7	Sequence 7, Appl
36	51	25.0	2973	2	US-08-821-355A-7	Sequence 7, Appl
37	51	25.0	2973	2	US-09-003-687A-7	Sequence 7, Appl
38	51	25.0	2973	4	US-09-136-605-7	Sequence 7, Appl
39	49.5	24.3	248	1	US-08-313-553-3	Sequence 3, Appl
40	49.5	24.3	248	1	US-08-303-651-2	Sequence 2, Appl
41	49.5	24.3	248	3	US-08-767-993-3	Sequence 3, Appl
42	49.5	24.3	262	1	US-08-313-553-2	Sequence 2, Appl
43	49.5	24.3	262	3	US-08-767-993-2	Sequence 2, Appl
44	49.5	24.3	559	1	US-08-313-553-15	Sequence 15, Appl
45	49.5	24.3	559	3	US-08-767-993-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-08-982-785A-10
; Sequence 10, Application US/08982785A
; Patent No. 6258929
; GENERAL INFORMATION:
; APPLICANT: Kosik, Kenneth S.
; APPLICANT: Zhou, Jianhua
; TITLE OF INVENTION: ALARM RELATED PEPTIDES AND
; NUCLEIC ACIDS AND DIAGNOSIS USING THEM
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/982,785A
; FILING DATE: 02-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/031,556
; FILING DATE: 02-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 05311/018001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 620 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-982-785A-10

Query Match 26.2%; Score 53.5; DB 4; Length 620;
Best Local Similarity 40.0%; Pred. No. 9.4;
Matches 14; Conservative 6; Mismatches 14; Indels 1;
Gaps 1;
QY 6 LAIFKSGGIPALVKMLGSPVDSVLFYAIT-TLHNL 39

STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/933,227
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0394 US
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 529 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1708480
US-08-933-227-4

Query Match 25.5%; Score 52; DB 2; Length 529;
Best Local Similarity 33.3%; Pred. No. 13;
Matches 11; Conservative 6; Mismatches 16; Indels 0; Gaps 0;

QY 7 AIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 39
: : | | | | : : | | | : : : : :
Db 157 AVDGGAIAPISLASHPAHISEQAVWALGNI 189

RESULT 8
US-08-982-785A-8
Sequence 8, Application US/08982785A
Patent No. 6258929
GENERAL INFORMATION:
APPLICANT: Kosik, Kenneth S.
APPLICANT: Zhou, Jianhua
TITLE OF INVENTION: ALARM RELATED PEPTIDES AND
TITLE OF INVENTION: NUCLEIC ACIDS AND DIAGNOSIS USING THEM
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/982,785A
FILING DATE: 02-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/031,556
FILING DATE: 02-DEC-1996
ATTORNEY/AGENT INFORMATION:

NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 05311/018001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 686 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-982-785A-8

Query Match 25.2%; Score 51.5; DB 4; Length 686;
Best Local Similarity 40.0%; Pred. No. 21;
Matches 14; Conservative 6; Mismatches 14; Indels 1; Gaps 1;

QY 6 LAIFKSGGIPALVKMLGSPVDSVLFYAIT-TLHNL 39
: : | | | | | : : : : :
Db 193 IALKNGGIPALVRLRLKTTDLRELVTVGLWNL 227

RESULT 9
US-08-982-785A-2
Sequence 2, Application US/08982785A
Patent No. 6258929
GENERAL INFORMATION:
APPLICANT: Kosik, Kenneth S.
APPLICANT: Zhou, Jianhua
TITLE OF INVENTION: ALARM RELATED PEPTIDES AND
TITLE OF INVENTION: NUCLEIC ACIDS AND DIAGNOSIS USING THEM
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/982,785A
FILING DATE: 02-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/031,556
FILING DATE: 02-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 05311/018001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 756 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-982-785A-2

Query Match 25.2%; Score 51.5; DB 4; Length 756;
Best Local Similarity 40.0%; Pred. No. 24;


```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: US/08/484,223B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-029
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
;
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 856 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
;
; US-08-484-223B-103

```

Query Match 25.0%; Score 51; DB 3; Length 856;
Best Local Similarity 33.3%; Pred. No. 33;
Matches 15; Conservative 6; Mismatches 16; Indels

Qy 1 HREGLLAIFKSGGIPALVKMLG--SPVDSVLFY-----AITTLH 37
| |||| :|| || | : || | : | |
Db 77 HTEGLLMVFKDNIIPYSFKVRSYTKIVTNILIIYNGWYADSVNRRH 121

RESULT 13
US-08-919-597-103
; Sequence 103, Application US/08919597
: Patent No. 6054265

? PATENT NO.: 0034283
 ? GENERAL INFORMATION:
 ?
 ? APPLICANT: Bolognesi, Dani P.
 ? APPLICANT: Matthews, Thomas J.
 ? APPLICANT: Wild, Carl T.
 ? APPLICANT: Barney, Shawn O.
 ? APPLICANT: Lambert, Dennis M.
 ? APPLICANT: Petteway, Stephen R.
 ? APPLICANT: Langlois, Alphonse J.
 ? TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
 ? OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
 ?
 ? TITLE OF INVENTION: TRANSMISSION
 ?
 ? NUMBER OF SEQUENCES: 273
 ? CORRESPONDENCE ADDRESS:
 ?
 ?

ADDRESS: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

: CLASSIFICATION:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/470,896
 : FILING DATE: 06-JUN-1995
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Coruzzi, Laura A.
 : REGISTRATION NUMBER: 30,742
 : REFERENCE/DOCKET NUMBER: 78772-020
 : TELECOMMUNICATION INFORMATION:
 :

```

; TELEPHONE: (212) 790-9090
;
; TELEFAX: (212) 869-9741/8864
;
; TELEX: 66141 PENNIE
;
; INFORMATION FOR SEQ ID NO: 103:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 856 amino acids
;
; TYPE: amino acid
;
; STRANDEDNESS: 1
;
; TOPOLOGY: unknown
;
; MOLECULE TYPE: protein
;
; US-08-919-597-103

```

Query Match 25.0%; Score 51; DB 3; Length 856;
Best Local Similarity 33.3%; Pred. No. 33;
Matches 15; Conservative 6; Mismatches 16; Indels

QY : 1' HREGLLAIFKSGGIPALVKMLG--SPVDSVLFY-----AITTLH 37
| | | | | : | | | | : | | | | : | | | |
Db 77 HTEGLLMVEKDNIIPYSFKVRSYTKIVTNILIIYNGWYADSVNRR 121

RESULT 14'
US-08-475-668A-103
; Sequence 103, Application US/08475668A
. Patent No. 6060065

8; Gaps 2;

Qy 1 HREGLLAIFKSGIPALVKMLG--SPVDSVLFY-----AITTLH 37
| | | | | : | | | | : | | | : | |
Db 77 HTEGLLMVFKDNIIPYSPKVRSTKIVTNILLYNGWYADSVNRRH 121

```

NUMBER OF SEQUENCES: 211
CONFERENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,668A
FILING DATE: 07-JUN-1995

```

CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-026

TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE

```

> INDEX: P06898
> INFORMATION FOR SEQ ID NO: 103:
>     LENGTH: 856 amino acids
>     TYPE: amino acid
>     STRANDEDNESS:
>     TOPOLOGY: unknown
>     MOLECULE TYPE: protein
>     US-08-475,668A-103

```

Query Match : 25.0%; Score 51; DB 3; Length 856;
Best Local Similarity 33.3%; Pred. No. 33;
Matches 0.15: Conservative 6; Mismatches 16; Indels

QY 01 HREGLLAIFKSGGIPALVKMLG--SPVDSVLFY-----AITTLH 37

DB 77 HTEGLLMVFKDNIIPYSFKVRSYTKIVTNILYNGWYADSVNHRH 121

RESULT 15

US-08-485-551A-103
; Sequence 103, Application US/08485551A
; Patent No. 6068973

GENERAL INFORMATION:

; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 211
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/485,551A
; FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

; NAME: Coruzzi, Laura A.

; REGISTRATION NUMBER: 30,742

; REFERENCE/DOCKET NUMBER: 7872-023

TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 790-9090

; TELEFAX: (212) 869-9741/8864

; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 103:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 856 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: unknown

; MOLECULE TYPE: protein

US-08-485-551A-103

Query Match 25.0%; Score 51; DB 3; Length 856;

Best Local Similarity 33.3%; Pred. No. 33;

Matches 15; Conservative 6; Mismatches 16; Indels 8; Gaps 2;

QY 1 HREGGLAIFKSGIPALVKMLG--SPVDSVLFY-----AITLH 37

DB 77 HTEGLLMVFKDNIIPYSFKVRSYTKIVTNILYNGWYADSVNHRH 121

Search completed: July 29, 2002, 16:05:07

Job time: 477 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 15:57:15 ; Search time 34.24 Seconds
(without alignments)
46.364 Million cell updates/sec

Title: US-09-641-104A-6
Perfect score: 204
Sequence: 1 HREGLAIPKSGGIPALVKM.....GSPVDSVLFYAITTLNLLL 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	204	100.0	781	1 CTNB_HUMAN	P35222 homo sapien
2	204	100.0	781	1 CTNB_MOUSE	Q02248 mus musculus
3	204	100.0	781	1 CTNB_RAT	Q9wu82 rattus norv
4	204	100.0	781	1 CTNB_XENLA	P26233 xenopus lae
5	191	93.6	621	1 PLAK_MOUSE	Q02257 mus musculus
6	191	93.6	743	1 PLAK_HUMAN	P14923 homo sapien
7	188	92.2	738	1 PLAK_XENLA	P30998 xenopus lae
8	188	92.2	813	1 ARM_MUSDO	Q02453 musca domes
9	188	92.2	843	1 ARM_DROME	P18824 drosophila
10	184	90.2	818	1 CTNB_URECA	P35224 urechis cau
11	179	87.7	820	1 CTNB_TRIGR	P35223 tripeustes
12	64	31.4	531	1 INA2_ARATH	O04294 arabidopsis
13	58	28.4	578	1 VAC8_MOUSE	P39968 saccharomyc
14	56	27.5	465	1 YB57_YEAST	P38310 saccharomyc
15	56	27.5	527	1 INAL_YCES	O22478 lycopersico
16	54.5	26.7	962	1 ARVC_HUMAN	O00192 homo sapien
17	54.5	26.7	1211	1 PKP4_HUMAN	Q95569 homo sapien
18	54	26.5	969	1 ARVC_MOUSE	P38203 mus musculus
19	53.5	26.2	911	1 CTDL_MOUSE	P30999 mus musculus
20	53.5	26.2	968	1 CTDL_HUMAN	O60716 homo sapien
21	53	26.0	433	1 FTSL_ARATH	Q42544 arabidopsis
22	52.5	25.7	311	1 YXFE_BAGSU	Q07835 bacillus su
23	52.5	25.7	350	1 EGSA_SULTO	P58460 sulfolobus
24	52.5	25.7	351	1 Y887_METH	O26973 methanobact
25	52	25.5	372	1 CNTR_RAT	Q08406 rattus norv
26	52	25.5	521	1 INA3_HUMAN	O00505 homo sapien
27	52	25.5	521	1 INA3_MOUSE	O35344 mus musculus
28	52	25.5	521	1 INA4_HUMAN	O00629 homo sapien
29	52	25.5	521	1 INA4_MOUSE	O35343 mus musculus
30	52	25.5	529	1 INA2_HUMAN	P52292 homo sapien
31	52	25.5	529	1 INA2_MOUSE	P52293 mus musculus
32	52	25.5	934	1 CLTC_HUMAN	P11586 h c-l-tetra
33	51	25.0	358	1 HMD_METJA	Q58194 methanococc

RESULT 1

ID	CTNB_HUMAN	STANDARD	PRT	781 AA
AC	P35222			
DT	01-FEB-1994 (Rel. 28, Created)			
DT	01-FEB-1994 (Rel. 28, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Beta-catenin.			
GN	CTNBN1 OR CTNNB1			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Placenta;			
RA	MEDLINE=95105247; PubMed=7806582;			
RT	Huelsen J., Birchmeier W., Behrens J.;			
RT	"E-cadherin and APC compete for the interaction with beta-catenin and the cytoskeleton.";			
RL	J. Cell Biol. 127:2061-2069(1994).			
RN	[2]			
RP	REVIEW			
RX	MEDLINE=20145417; PubMed=10679188;			
RA	Kikuchi A.;			
RT	"Regulation of beta-catenin signaling in the Wnt pathway.";			
RL	Biochem. Biophys. Res. Commun. 268:243-248(2000).			
RN	[3]			
RP	X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 133-664.			
RX	MEDLINE=20578806; PubMed=11136974;			
RA	Graham T.A., Weaver C., Mao F., Kimmelman D., Xu W.;			
RT	"Crystal structure of a beta-catenin/Tcf complex.";			
RL	Cell 103:885-896(2000).			
CC	1- FUNCTION: INVOLVED IN THE REGULATION OF CELL ADHESION AND IN SIGNAL TRANSDUCTION THROUGH THE WNT PATHWAY.			
CC	1- SUBUNIT: THE CYTOSOLIC FORM BOUND TO AXIN ALLOWS THE STIMULATION OF ITS DEGRADATION. THE CYTOSOLIC FORM ALSO INTERACTS WITH APC AND E-CADHERIN. THE NUCLEAR FORM BINDS TCF/LEF-1 AND MAY ALSO BIND PONTIN52 AND DUPLIN. CAN HETERODIMERIZE WITH ALPHA-CATENIN AND ALSO FORM A TRIMERIC COMPLEX WITH CADHERIN. MEDIATES THE ANCHORAGE OF THE CADHERIN COMPLEX TO THE CORTICAL ACTIN CYTOSKELETON.			
CC	1- SUBCELLULAR LOCATION: CYTOPLASMIC WHEN IT IS UNSTABILIZED (HIGH LEVEL OF PHOSPHORYLATION). TRANSLOCATES TO THE NUCLEUS WHEN IT IS STABILIZED (LOW LEVEL OF PHOSPHORYLATION).			
CC	1- PTM: PROBABLY PHOSPHORYLATED BY GSK-3B. WNT-DEPENDENT ACTIVATION OF DVL ANTAGONIZES THE ACTION OF GSK-3B. PHOSPHORYLATION OF BETA-CATENIN IS REDUCED AND THE PROTEIN IS NO LONGER DEGRADED, RESULTING IN ITS ACCUMULATION IN CYTOPLASM.			
CC	1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.			
CC	1- SIMILARITY: CONTAINS 12 ARM REPEATS.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use, by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial			

005680 mycobacteri
014063 schizosacch
Q57689 methanococc
P03188 epstein-bar
P45128 haemophilus
P70478 rattus norv
P25054 homo sapien
Q61315 mus musculu
Q60396 apodemus sy
P09368 saccharomyc
P33073 peptostrept
P09180 drosophila

16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
01-MAR-2002 (Rel. 41, Last annotation update)
Beta-catenin.
CTNNB1.
Rattus norvegicus (Rat).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE=Testis;
RX MEDLINE=99428593; PubMed=10497305;
RA Chung S.S.W., Lee W.M., Cheng C.Y.;
RT "Study on the formation of specialized inter-Sertoli cell junctions in
vitro.";
RL J. Cell. Physiol. 181:258-272(1999).
CC -!- FUNCTION: INVOLVED IN THE REGULATION OF CELL ADHESION AND IN
CC SIGNAL TRANSDUCTION THROUGH THE WNT PATHWAY (BY SIMILARITY).
CC -!- SUBUNIT: THE CYTOSOLIC FORM BOUND TO AXIN ALLOWS THE STIMULATION
CC OF ITS DEGRADATION. THE CYTOSOLIC FORM ALSO INTERACTS WITH APC AND
CC E-CADHERIN. THE NUCLEAR FORM BINDS TCF/LEF-1 AND MAY ALSO BIND
CC PONTIN52 AND DUPLIN. CAN HETERODIMERIZE WITH ALPHA-CATENIN AND
CC ALSO FORM A TRIMERIC COMPLEX WITH CADHERIN. MEDIATES THE ANCHORAGE
CC OF THE CADHERIN COMPLEX TO THE CORTICAL ACTIN CYTOSKELETON (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC WHEN IT IS UNSTABILIZED (HIGH
CC LEVEL OF PHOSPHORYLATION). TRANSLOCATES TO THE NUCLEUS WHEN IT IS
CC STABILIZED (LOW LEVEL OF PHOSPHORYLATION) (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE TESTIS.
CC -!- DEVELOPMENTAL STAGE: HIGHLY EXPRESSED AT E30-60 DAY DPC IN THE
CC TESTIS. REDUCED EXPRESSION AT E90 DAY DPC.
CC -!- PTM: PROBABLY PHOSPHORYLATED BY GSK-3B. WNT-DEPENDENT ACTIVATION
CC OF DVL ANTAGONIZES THE ACTION OF GSK-3B. PHOSPHORYLATION OF BETA-
CC CATENIN IS REDUCED AND THE PROTEIN IS NO LONGER DEGRADED.
CC RESULTING IN ITS ACCUMULATION IN CYTOPLASM (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -!- SIMILARITY: CONTAINS 12 ARM REPEATS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).

EMBL; AF121265; AAD28504.1; -
DR HSSP; Q02248; 2BCT.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 12.
DR SMART; SM00185; ARM; 11.
DR PROSITE; PS50176; ARM_REPEAT; 9.
KW Cell adhesion; Cytoskeleton; Structural protein; Nuclear protein;
KW Repeat.
FT REPEAT 151 191 ARM 1.
FT REPEAT 193 234 ARM 2.
FT REPEAT 235 276 ARM 3.
FT REPEAT 277 318 ARM 4.
FT REPEAT 319 360 ARM 5.
FT REPEAT 361 389 ARM 6.
FT REPEAT 400 441 ARM 7.
FT REPEAT 442 484 ARM 8.
FT REPEAT 489 530 ARM 9.
FT REPEAT 531 571 ARM 10.
FT REPEAT 594 636 ARM 11.
FT REPEAT 637 666 ARM 12.
SQ SEQUENCE 781 AA; 85454 MW; 9C29186B6D54B87 CRC64;

Query Match 100.0%; Score 204; DB 1; Length 781;
Best Local Similarity 100.0%; Pred. No. 1.3e-18;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 HREGLLAIFKSGIPALVKMLGSPVDSVLFYAITTLNLLL 41
24 HREGLLAIFKSGIPALVKMLGSPVDSVLFYAITTLNLLL 264
RESULT 4
CTNB_XENLA
ID CTNB_XENLA STANDARD; PRT; 781 AA.
AC P26233;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Beta-catenin.
OS Xenopus laevis, (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RC SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=92073903; PubMed=1962194;
RA McCrea P.D., Turk C.W., Gumbiner B.M.;
RT "A homolog of the armadillo protein in Drosophila (plakoglobin)
associated with E-cadherin.";
RL Science 254:1359-1361(1991).
CC -!- FUNCTION: BINDS TO THE CYTOPLASMIC DOMAIN OF THE CELL-CELL
CC ADHESION MOLECULE E-CADHERIN, AND PERHAPS TO OTHER (MEMBRANE)
CC PROTEINS. THE ASSOCIATION OF CATENINS TO CADHERINS PRODUCES A
CC COMPLEX WHICH IS LINKED TO THE ACTIN FILAMENT NETWORK, AND WHICH
CC SEEMS TO BE OF PRIMARY IMPORTANCE FOR CADHERINS CELL-ADHESION
CC PROPERTIES. ESSENTIAL COMPONENT OF THE VERTEBRATE WNT-SIGNALING
CC PATHWAY AND IS REQUIRED FOR FORMATION OF THE BETA-CATENIN FAMILY.
CC -!- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -!- SIMILARITY: CONTAINS 10 ARM REPEATS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).

EMBL; M77013; AAA49670.1; -
DR PIR; S35099; S35099.
DR HSSP; Q02248; 2BCT.
DR TRANSFAC; T03026; -
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 12.
DR SMART; SM00185; ARM; 11.
DR PROSITE; PS50176; ARM_REPEAT; 9.
KW Cell adhesion; Cytoskeleton; Structural protein; Repeat.
KW Repeat.
FT REPEAT 141 180 ARM 1.
FT REPEAT 181 225 ARM 2.
FT REPEAT 226 264 ARM 3.
FT REPEAT 267 306 ARM 4.
FT REPEAT 307 350 ARM 5.
FT REPEAT 351 391 ARM 6.
FT REPEAT 392 432 ARM 7.
FT REPEAT 433 473 ARM 8.
FT REPEAT 474 519 ARM 9.
FT REPEAT 520 562 ARM 10.
FT REPEAT 563 604 ARM 11.
FT REPEAT 605 646 ARM 12.
SQ SEQUENCE 781 AA; 85449 MW; 3ECD27232239F799 CRC64;

Query Match 100.0%; Score 204; DB 1; Length 781;
Best Local Similarity 100.0%; Pred. No. 1.3e-18;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 HREGLLAIFKSGIPALVKMLGSPVDSVLFYAITTLNLLL 41
24 HREGLLAIFKSGIPALVKMLGSPVDSVLFYAITTLNLLL 264

Db 224 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHLL 264

```

RESULT 5
PLAK_MOUSE STANDARD; PRT; 621 AA.
AC Q02257;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Junction plakoglobin (Desmoplakin III) (fragment).
GN JUP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92376536; PubMed=1509266;
RA Butz S., Steppert J., Weissig H., Kemler R.;
RT "Plakoglobin and beta-catenin: distinct but closely related.";
RL Science 257:1142-1144(1992).
RN [2]
RP REVISIONS TO 294 AND 296.
RA Butz S.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: COMMON JUNCTIONAL PLAQUE PROTEIN. THE MEMBRANE-
CC ASSOCIATED PLAQUES ARE ARCHITECTURAL ELEMENTS IN AN IMPORTANT
CC STRATEGIC POSITION TO INFLUENCE THE ARRANGEMENT AND FUNCTION
CC OF BOTH THE CYTOSKELETON AND THE CELLS WITHIN THE TISSUE. THE
CC PRESENCE OF PLAKOGLOBIN IN BOTH THE DESMOSOMES AND IN THE
CC INTERMEDIATE JUNCTIONS SUGGESTS THAT IT PLAYS A CENTRAL ROLE IN
CC THE STRUCTURE AND FUNCTION OF SUBMEMBRANOUS PLAQUES.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC IN A SOLUBLE AND MEMBRANE-
CC ASSOCIATED FORM.
CC -1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -1- SIMILARITY: CONTAINS AT LEAST 9 ARM REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M90365; AAB02885.1; -
CC PIR; S35092; S35092.
CC HSSP; Q02248; 1DOW.
CC DR MGD; MGI:96650; Jup.
CC DR InterPro; IPR000225; Armadillo.
CC DR Pfam; PF00514; Armadillo_seg; 11.
CC DR SMART; SM00185; ARM; 8.
CC DR PROSITE; PS50176; ARM_REPEAT; 9.
CC KW Cell adhesion; Cytoskeleton; Structural protein; Repeat.
CC FT NON_TER 1 1
CC FT REPEAT 8 47 ARM 1.
CC FT REPEAT 92 131 ARM 2.
CC FT REPEAT 134 173 ARM 3.
CC FT REPEAT 218 257 ARM 4.
CC FT REPEAT 259 296 ARM 5.
CC FT REPEAT 299 340 ARM 6.
CC FT REPEAT 346 386 ARM 7.
CC FT REPEAT 388 427 ARM 8.
CC FT REPEAT 450 489 ARM 9.
CC SEQUENCE 621 AA; 68111 MW; 17CF444607422BAA CRC64;

```

Query Match 93.6%; Score 191; DB 1; Length 621;
 Best Local Similarity 92.7%; Pred. No. 4.7e-17;
 Matches 38; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHLL 41
 Db 91 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHLL 131

```

RESULT 6
PLAK_HUMAN STANDARD; PRT; 743 AA.
AC P14923;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Junction plakoglobin (Desmoplakin III).
GN JUP OR DP3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89264555; PubMed=2726765;
RA Franke W.W., Goldschmidt M.D., Zimbelmann R., Mueller H.M.,
RA Schiller D.L., Cowin P.;
RT "Molecular cloning and amino acid sequence of human plakoglobin, the
RT common junctional plaque protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:4027-4031(1989).
CC -1- FUNCTION: COMMON JUNCTIONAL PLAQUE PROTEIN. THE MEMBRANE-
CC ASSOCIATED PLAQUES ARE ARCHITECTURAL ELEMENTS IN AN IMPORTANT
CC STRATEGIC POSITION TO INFLUENCE THE ARRANGEMENT AND FUNCTION
CC OF BOTH THE CYTOSKELETON AND THE CELLS WITHIN THE TISSUE. THE
CC PRESENCE OF PLAKOGLOBIN IN BOTH THE DESMOSOMES AND IN THE
CC INTERMEDIATE JUNCTIONS SUGGESTS THAT IT PLAYS A CENTRAL ROLE IN
CC THE STRUCTURE AND FUNCTION OF SUBMEMBRANOUS PLAQUES.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC IN A SOLUBLE AND MEMBRANE-
CC ASSOCIATED FORM.
CC -1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -1- SIMILARITY: CONTAINS 8 ARM REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M23410; AAA64895.1; -
CC PIR; A32905; A32905.
CC HSSP; Q02248; 2BCT.
CC DR MIM; MIM:173325; -
CC DR InterPro; IPR000225; Armadillo.
CC DR Pfam; PF00514; Armadillo_seg; 10.
CC DR SMART; SM00185; ARM; 8.
CC DR PROSITE; PS50176; ARM_REPEAT; 8.
CC KW Cell adhesion; Cytoskeleton; Structural protein; Repeat.
CC FT INIT_MET 0 0 PROBABLE.
CC FT REPEAT 141 179 ARM 1.
CC FT REPEAT 183 224 ARM 2.
CC FT REPEAT 225 252 ARM 3.
CC FT REPEAT 308 351 ARM 4.
CC FT REPEAT 389 430 ARM 5.
CC FT REPEAT 431 473 ARM 6.
CC FT REPEAT 478 521 ARM 7.
CC FT REPEAT 582 624 ARM 8.
CC FT CONFLICT 96 139 GOLA -> DSSL (IN REF. 1; AAA64895).
CC FT CONFLICT 139 139 V -> A (IN REF. 1; AAA64895).
CC SEQUENCE 743 AA; 81498 MW; 472741F400D388FD CRC64;

```

Query Match 93.6%; Score 191; DB 1; Length 743;
 Best Local Similarity 92.7%; Pred. No. 5.7e-17;
 Matches 38; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Tue Jul 30 08:15:37 2002

RA REPEAT 587 626 ARM 9.
 RA REPEAT 628 667 ARM 10.
 SQ SEQUENCE 813 AA; 88237 MW; 73E61F59BDBFA580 CRC64;

 Query Match 92.2%; Score 188; DB 1; Length 813;
 Best Local Similarity 90.2%; Pred. No. 1.5e-16;
 Matches 37; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

 QY 1 HREGLLAIFKSGGIPALVKMLGSPVDSVLFVAITLHLL 41
 DB 222 HQGLLAIFKSGGIPALVKMLGSPVDSVLFVAITLHLL 262

 RESULT 9
 ARM_DROME STANDARD; PRT; 843 AA.
 ID ARM_DROME Q9W546;
 AC P18824; 002371; 09W546;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Armadillo segment polarity protein.
 GN ARM OR EG:86E4.6 OR CG11579.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A. (CYTOPLASMIC ISOFORM).
 RC STRAIN=OREGON-R;
 RX MEDLINE=89211895; PubMed=2707602;
 RA Riggleman B., Wieschaus E., Schedl P.;
 RT "Molecular analysis of the armadillo locus: uniformly distributed
 transcripts and a protein with novel internal repeats are associated
 with a Drosophila segment polarity gene.";
 RL Genes Dev. 3:96-113(1989).
 RN [2]
 RP SEQUENCE FROM N.A. FUNCTION, TISSUE SPECIFICITY, AND ALTERNATIVE
 SPlicing.
 RC TISSUE=Head;
 RX MEDLINE=98298928; PubMed=9635189;
 RA Loureiro J., Peifer M.;
 RT "Roles of Armadillo, a Drosophila catenin, during central nervous
 system development.";
 RL Curr. Biol. 8:622-632(1998).
 RN [3]
 RP SEQUENCE FROM N.A. (CYTOPLASMIC ISOFORM).
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer V.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Fertlira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,

Merklov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 Palazkolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 Glods R.A., Myers E.W., Rubin G.M., Venter J.C.;
 "The genome sequence of Drosophila melanogaster.";
 Science 287:2185-2195(2000).
 [4]
 RP SEQUENCE FROM N.A. (CYTOPLASMIC ISOFORM).
 RC STRAIN=OREGON-R;
 RX MEDLINE=20196011; PubMed=10731137;
 RA Benos P.V., Gatt M.K., Ashburner M., Murphy L., Harris D.,
 Barrell B.G., Ferraz C., Vidal S., Brun C., Demallies J., Cadieu E.,
 Dreano S., Gloux S., Leilaure V., Mottier S., Gallibert F., Borkova D.,
 Minana B., Kafatos F.C., Louis C., Siden-Kiamos I., Bolshakov S.,
 Papagiannakis G., Spanos L., Cox S., Madueno E., de Pablos B.,
 Moudell J., Peter A., Schoettler P., Werner M., Mourikioti F.,
 Beinert N., Dowe G., Schaefer U., Jaekle H., Bucheton A.,
 Callister D.M., Campbell L.A., Darlamsou A., Henderson N.S.D.,
 McMillan P.J., Salles C., Tait E.A., Valenti P., Saunders R.D.C.,
 Glover D.M.;
 "From sequence to chromosome: the tip of the X chromosome of D.
 melanogaster.";
 Science 287:2220-2222(2000).
 [5]
 RP PHOSPHORYLATION.
 RX MEDLINE=95113174; PubMed=7529201;
 RA Peifer M., Pai L.-M., Casey M.;
 RT "Phosphorylation of the Drosophila adherens junction protein
 Armadillo: roles for wingless signal and zeste-white 3 kinase.";
 RL Dev. Biol. 166:543-556(1994).
 CC -1- FUNCTION: NEURAL ISOFORM MAY ASSOCIATE WITH CADN AND PARTICIPATE
 IN THE TRANSMISSION OF DEVELOPMENTAL INFORMATION. CAN ASSOCIATE
 WITH ALPHA-CATENIN. CYTOPLASMIC ISOFORM ACCUMULATES THROUGH WG
 SIGNALING. ARM-FUNCTION IN WG SIGNAL TRANSDUCTION IS REQUIRED
 EARLY IN DEVELOPMENT FOR DETERMINATION OF NEUROBLAST FATE. ARM AND
 ABL PROTEINS FUNCTION COOPERATIVELY AT ADHERENS JUNCTIONS IN BOTH
 THE CNS AND EPIDERMIS.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC AND MEMBRANE-ASSOCIATED WITH THE
 INNER SURFACE OF CELL MEMBRANE.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; CYTOPLASMIC (SHOWN HERE) AND
 NEURAL; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: CYTOPLASMIC ISOFORM IS PREDOMINANT BEFORE GERM
 BAND RETRACTION. AFTER RETRACTION AND DURING LARVAL STAGES, IT IS
 FOUND IN HIGH LEVELS IN SPECIFIC CELLS ALONG THE CNS MIDLINE.
 CC -1- NEURAL ISOFORM IS FIRST SEEN AFTER GERM BAND RETRACTION IN THE
 AXON TRACTS OF THE CNS, ALSO PRESENT IN AXONS DURING LARVAL STAGES
 AND ACCUMULATES IN THE MOTOR NEURONS OF THE SEGMENTAL AND
 INTERSEGMENTAL NERVES AS THEY EXIT THE CNS. BOTH ISOFORMS
 ACCUMULATE IN THE PNS.
 CC -1- DEVELOPMENTAL STAGE: PRESENT AT ALL STAGES, BUT REACHED THE
 HIGHEST LEVELS DURING EARLY TO MID-EMBRYOGENESIS.
 CC -1- PTM: PHOSPHORYLATED ON SER, THR AND TYR RESIDUES. LEVEL OF
 PHOSPHORYLATION VARIES BOTH DURING EMBRYONIC DEVELOPMENT AND FROM
 EMBRYONIC TISSUE TO TISSUE. SGG IS REQUIRED FOR PHOSPHORYLATION.
 CC -1- AND WG SIGNAL NEGATIVELY REGULATES ARM PHOSPHORYLATION.
 CC -1- HYPOPHOSPHORYLATED FORM OF ARM INCREASES IN STEADY-STATE LEVELS.
 CC -1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 12.5 ARM REPEATS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; L10354; AAA30089.1; -;
 DR PIR; S33794; S33794.
 DR HSSP; Q02248; LDOW.
 DR InterPro; IPR000225; Armadillo.
 DR Pfam; PF00514; Armadillo_seg; 12.
 DR SMART; SM00185; ARM; 12.
 DR PROSITE; PS0176; ARM_REPEAT; 9.
 KW Cell adhesion; Cytoskeleton; Structural protein; Repeat.
 FT REPEAT 157 196
 FT REPEAT 139 239
 FT REPEAT 241 280
 FT REPEAT 283 322
 FT REPEAT 367 405
 FT REPEAT 406 445
 FT REPEAT 448 489
 FT REPEAT 495 535
 FT REPEAT 603 642
 FT REPEAT 644 683
 FT SEQUENCE 820 AA; 89361 MW; 57235E0F57795FD3 CRC64;
 SQ
 Query Match 87.7%; Score 179; DB 1; Length 820;
 Best Local Similarity 85.4%; Pred. No. 2.2e-15;
 Matches 35; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 1 HREGLLAIFKGGIPALVLMGSPVDSVLFYAITTLHLL 41
 DB 240 HRAGLLIQFSGGIPALIKLSSPVESVLFYAITTLHLL 280
 RESULT 12
 ID IMA2_ARATH STANDARD; PRT; 531 AA.
 AC Q04294; Q04255;
 DT 13-JUL-1998 (Rel. 36, Created)
 DT 13-JUL-1998 (Rel. 36, Last sequence update)
 DE Importin alpha-2 subunit (Karyopherin alpha-2 subunit) (KAP alpha).
 GN A74G02150 OR T10M13.16.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Salchert K.D.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=20083488; PubMed=10617198;
 RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
 RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
 RA Harris B., Ansorge W., Brandt P., Grivell L.A., Rieger M.,
 RA Weichselgartner M., de Simone V., Obermaier L.B., Mache R., Mueller M.,
 RA Kreis M., Delsen M., Puigdomenech P., Watson M., Schmidtheini T.,
 RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
 RA Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
 RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
 RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
 RA Braeken M., Welltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
 RA Weitzenecker T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
 RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
 RA Moolman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
 RA Berneiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
 RA De Keyser A., Buysschaert C., Gielen J., Villarroel R., De Clercq R.,
 RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
 Clark L., Doggett J., Hall S., Kay M., Lennard N., McIlroy K., Mayes R.,
 Pételt A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
 Borkova D., Blocker H., Scharfe M., Grimm M., Loehner T.-H.,
 Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
 Gabel C., Fuchs M., Farmann B., Granderath K., Bauner D., Herzl A.,
 Neumann S., Argirou A., Vitale D., Liguori R., Piravandi E.,
 Massenet O., Quigley F., Clabault G., Muendlein A., Felber R.,
 Schöhl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 Chaffor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
 Gibbons T., Weber N., Vandenbol M., Barques M., Terol J., Torres A.,
 Pérez-Pérez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
 Heijhen L., Schwarz S., Scholler P., Heber S., Stocker S.,
 Frishman D., Haase D., Lemcke K., Mewes H.-W., Habermann K.,
 Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Spiegel L.,
 Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Hapler L.,
 Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
 Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
 Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K., Hillier L.,
 Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Layman D.,
 Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
 Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
 Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
 Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
 Svaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
 Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
 Chen E., Marra M., Martienssen R., McCombie W.R.;
 "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 thaliana".
 RL Nature 402:769-777 (1999).
 CC -!- FUNCTION: BINDS SPECIFICALLY AND DIRECTLY TO SUBSTRATES CONTAINING
 CC EITHER A SIMPLE OR BIPARTITE NLS MOTIF. PROMOTES DOCKING OF IMPORT
 CC SUBSTRATES TO THE NUCLEAR ENVELOPE. SEEMS TO ACT AS A CYTOSOLIC
 CC RECEPTOR FOR BOTH SIMPLE AND BIPARTITE NLS MOTIFS (BY SIMILARITY).
 CC -!- SUBUNIT: FORMS A COMPLEX WITH IMPORTIN BETA-1 SUBUNIT (BY
 CC SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE IMPORTIN ALPHA FAMILY.
 CC -!- SIMILARITY: CONTAINS 8 ARM REPEATS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

Tue Jul 30 08:15:37 2002

us-09-641-104a-6.rsp

```
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RA Kunik T., Mizrachi L., Citovsky V., Gafni Y.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: BINDS SPECIFICALLY AND DIRECTLY TO SUBSTRATES CONTAINING
CC EITHER A SIMPLE OR BIPARTITE NLS MOTIF. PROMOTES DOCKING OF IMPORT
CC SUBSTRATES TO THE NUCLEAR ENVELOPE. SEEMS TO ACT AS A CYTOSOLIC
CC RECEPTOR FOR BOTH SIMPLE AND BIPARTITE NLS MOTIFS (BY SIMILARITY).
CC -!- SUBUNIT: FORMS A COMPLEX WITH IMPORTIN BETA-1 SUBUNIT (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE IMPORTIN ALPHA FAMILY.
CC -!- SIMILARITY: CONTAINS 8 ARM REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF01752; AAC23722.1; -.
CC HSP; 002821; 1BK5.
CC InterPro: IPR000225; Armadillo.
CC InterPro: IPR002652; IBB.
CC Pfam; PF00514; Armadillo_seg; 8.
CC DR PF01749; IBB; 1.
CC DR SMART; SM00185; ARM; 8.
CC DR PROSITE; PS0176; ARM_REPEAT; 5.
CC Transport; Protein transport; Repeat.
FT DOMAIN 12 51
FT REPEAT 109 151 ARM 1.
FT REPEAT 152 196 ARM 2.
FT REPEAT 197 234 ARM 3.
FT REPEAT 235 279 ARM 4.
FT REPEAT 280 319 ARM 5.
FT REPEAT 320 362 ARM 6.
FT REPEAT 363 403 ARM 7.
FT REPEAT 404 445 ARM 8.
FT DOMAIN 446 527 ASP/GLU-RICH (ACIDIC)..
SQ SEQUENCE 527 AA; 58605 MW; 4A3F01691CE4817 CRC64;
```

```
Query Match 27.5%; Score 56; DB 1; Length 527;
Best Local Similarity 29.7%; Pred. No. 8.7;
Matches 11; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

QY 3 EGLIAIFKSGIPALVKMLGSPVDSVLFVAITTLHNL 39
   | : | : | : | : | : | : | : | : | : |
Db 150 ENTKVVIDYGSVPFIIFIRLLSSPDDVREQAVWALGNI 186
```

Search completed: July 29, 2002, 16:09:50
Job time: 755 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 16:04:10 ; Search time 124.4 seconds

(without alignments)

57.016 Million cell updates/sec

Title: US-09-641-104A-6

Perfect score: 204

Sequence: 1 HREGLLAIFKSGIPALVKM.....GSPVDSVLFYAITLHLL 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phase:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_virus:*

16: sp_bacteriap:*

17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	204	100.0	781	11 Q9D335	Q9D335 mus musculus
2	204	100.0	781	13 Q42486	Q42486 gallus gall
3	194	95.1	780	13 Q90424	Q90424 brachydanio
4	192	94.1	773	5 Q76152	Q76152 ciona savig
5	191	93.6	745	4 Q15151	Q15151 homo sapien
6	191	93.6	745	4 Q9BWC4	Q9BWC4 homo sapien
7	191	93.6	745	11 P70565	P70565 rattus norv
8	184	90.2	729	13 Q9PVF7	Q9PVF7 brachydanio
9	181	88.7	769	5 Q9NL44	Q9NL44 ciona intes
10	180	88.2	821	5 Q61229	Q61229 lytechinus
11	160	78.4	806	5 Q25100	Q25100 hydra magni
12	104	51.0	678	5 Q44326	Q44326 caenorhabdi
13	67	32.8	198	4 Q9H6L4	Q9H6L4 homo sapien
14	66	32.4	529	10 Q94KA9	Q94KA9 capsicum an
15	65	31.9	538	10 Q9FWY7	Q9FWY7 arabidopsis
16	64	31.4	531	10 Q49601	Q49601 arabidopsis

17	64	31.4	531	10 Q94F55	Q94F55 arabidopsis
18	63	30.9	180	10 Q9FHT6	Q9FHT6 arabidopsis
19	63	30.9	656	10 Q9FI13	Q9FI13 arabidopsis
20	62	30.4	534	10 Q9SLX0	Q9SLX0 oryza sativ
21	62	30.4	534	10 Q94GV3	Q94GV3 oryza sativ
22	61	29.9	442	5 Q9VRS0	Q9VRS0 drosophila
23	61	29.9	660	10 Q9SNC6	Q9SNC6 arabidopsis
24	59	28.9	233	8 Q9MP25	Q9MP25 anopheles s
25	59	28.9	272	8 Q9B293	Q9B293 ceratosolen
26	59	28.9	272	8 Q9B292	Q9B292 ceratosolen
27	59	28.9	639	10 Q9SV34	Q9SV34 arabidopsis
28	59	28.9	811	5 Q18825	Q18825 caenorhabdi
29	58	28.4	526	10 Q82783	Q82783 oryza sativ
30	58	28.4	548	3 Q9C2K9	Q9C2K9 neurospora
31	58	28.4	674	10 Q9L2W3	Q9L2W3 arabidopsis
32	58	28.4	698	10 Q9FN17	Q9FN17 arabidopsis
33	57	27.9	198	10 Q04286	Q04286 salaginella
34	57	27.9	264	10 Q94KD4	Q94KD4 arabidopsis
35	57	27.9	441	10 Q9FJ92	Q9FJ92 arabidopsis
36	57	27.9	528	10 Q49602	Q49602 arabidopsis
37	57	27.9	538	10 Q80480	Q80480 arabidopsis
38	57	27.9	779	5 Q9W5T9	Q9W5T9 drosophila
39	57	27.9	781	5 Q9NHP1	Q9NHP1 drosophila
40	56	27.5	425	16 Q9H0H7	Q9H0H7 pseudomonas
41	56	27.5	669	4 Q96M49	Q96M49 homo sapien
42	55	27.0	174	8 Q9G2V2	Q9G2V2 orcheseila
43	55	27.0	372	11 Q88507	Q88507 mus musculu
44	55	27.0	402	10 Q9S015	Q9S015 arabidopsis
45	55	27.0	404	10 Q94F85	Q94F85 oryza sativ

ALIGNMENTS

RESULT 1:

Q9D335 PRELIMINARY; PRT; 781 AA.

AC Q9D335;

DT 01-JUN-2001 (Tremblrel. 17, Created)

DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

DE ADULT MALE COLON CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,

DE CLONE:9030417H18, FULL INSERT SEQUENCE.

GN CATNB:

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

SEQUENCE FROM N.A.

RP STRAIN=C57BL/6J; TISSUE=COLON;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamakawa I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schrim L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gusticich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Watanabe-Borisi A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.,
RA "Functional annotation of a full-length mouse cDNA collection.",
RL Nature 409:685-690(2001).
DR EMEL; AK018515; BAB31250.1; -.

```
DR HSP; P35222; IG3J.
DR MGD; MGI:88276; Catnb.
DR InterPro; IPR00225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 12.
DR SMART; SM00185; ARM; 12.
DR PROSITE; PS50176; ARM_REPEAT; 7.
SQ SEQUENCE 781 AA; 85546 MW; 937538C3B5CD75D1 CRC64;

Query Match      100.0%; Score 204; DB 11; Length 781;
Best Local Similarity 100.0%; Pred. No. 1.1e-18;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 41
Db 224 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 264

RESULT 2
ID O42486 PRELIMINARY; PRT; 781 AA.
AC O42486;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE BETA CATENIN.
GN CHBCAT.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WHITE LEGHORN; TISSUE=DORSAL SKIN;
RX MEDLINE=97464068; PubMed=9322759;
RA Lu J., Chuong C.M., Wideltz R.B.;
RT "Isolation and characterization of chicken beta-catenin.";
RL Gene 196:201-207(1997).
DR EMBL; U82964; AAB80856.1; -.
DR HSP; P35222; IG3J.
DR InterPro; IPR00225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 12.
DR SMART; SM00185; ARM; 11.
DR PROSITE; PS50176; ARM_REPEAT; 8.
SQ SEQUENCE 781 AA; 85438 MW; 6D205D9A4DBAC562 CRC64;

Query Match      100.0%; Score 204; DB 13; Length 781;
Best Local Similarity 100.0%; Pred. No. 1.1e-18;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 41
Db 224 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 264

RESULT 3
ID Q90424 PRELIMINARY; PRT; 780 AA.
AC Q90424;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE B-CATENIN.
GN CTNNB.
OS Brachydanio rerio (zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=96122902; PubMed=8562427;
RT "Induction of a secondary embryonic axis in zebrafish occurs following the overexpression of beta-catenin.";
RL Mech. Dev. 53:261-273(1995).
DR EMBL; U41081; AAC59732.1; -.
DR HSP; P35222; IG3J.
DR ZFIN; ZDB-GENE-980526-362; cttnb.
DR InterPro; IPR00225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 12.
DR SMART; SM00185; ARM; 11.
DR PROSITE; PS50176; ARM_REPEAT; 8.
SQ SEQUENCE 780 AA; 85542 MW; D7A1FB90F94066DC CRC64;

Query Match      95.1%; Score 194; DB 13; Length 780;
Best Local Similarity 95.1%; Pred. No. 2.4e-17;
Matches 39; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 41
Db 223 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 263

RESULT 4
ID O76152 PRELIMINARY; PRT; 773 AA.
AC O76152;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE BETA-CATENIN.
OS Ciona savignyi.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;
OC Cionidae; Ciona.
OX NCBI_TaxID=51511;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98443204; PubMed=9769178;
RA Yoshioka S., Marikawa Y., Satoh N.;
RT "Regulation of the trunk-tail patterning in the ascidian embryo: a possible interaction of cascades between lithium/beta-catenin and localized maternal factor pem.";
RL Dev. Biol. 202:264-279(1998).
DR EMBL; AB012160; BAA32789.1; -.
DR HSP; P35222; IG3J.
DR InterPro; IPR00225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 11.
DR SMART; SM00185; ARM; 9.
DR PROSITE; PS50176; ARM_REPEAT; 7.
SQ SEQUENCE 773 AA; 85217 MW; C1340CF82AFEBDAB CRC64;

Query Match      94.1%; Score 192; DB 5; Length 773;
Best Local Similarity 87.8%; Pred. No. 4.3e-17;
Matches 36; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 41
Db 213 HKOGLLSIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 253

RESULT 5
ID Q15151 PRELIMINARY; PRT; 745 AA.
AC Q15151; Q15093;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PLAKOGLOBIN.
GN JUP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89264555; PubMed=2726765;
RA Franke W.W., Goldschmidt M.D., Zimbelmann R., Mueller H.M.,
RS Schiller D.L., Cowin P.;
RT "Molecular cloning and amino acid sequence of human plakoglobin, the
RL common junctional plaque protein."
RN Proc. Natl. Acad. Sci. U.S.A. 86:4027-4031(1989).
[2]
RP SEQUENCE FROM N.A.
RA Zimbelmann R.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 239-409 FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=96157724; PubMed=8576101;
RA Ozawa M., Nukui K., Toyoyama H., Ohi Y.;
RT "Cloning of an alternative form of plakoglobin (gamma-catenin) lacking
RL the fourth armadillo repeat."
RN J. Biochem. 118:836-840(1995).
[4]
RP SEQUENCE FROM N.A.
RX PubMed=11016852;
RA Whitlock N.V., Eady R.A.J., McGrath J.A.;
RT "Genomic Organization and Amplification of the human plakoglobin
RL gene."
RX Exp. Dermatol. 9:323-326(2000).
DR EMBL; Z68228; CAA92522.1; -
DR EMBL; D50808; BAA09435.1; -
DR EMBL; AF306723; AAG16727.1; -
DR EMBL; AF233882; AAG16727.1; JOINED.
DR HSP; Q02248; 2BCT.
DR InterPro: IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 11.
DR SMART; SM00185; ARM; 9.
DR PROSITE; PS0176; ARM_REPEAT; 7.
SQ SEQUENCE 745 AA; 81744 MW; 3519A0973748BCF4 CRC64;

Query Match 93.6%; Score 191; DB 4; Length 745;
Best Local Similarity 92.7%; Pred. No. 5.7e-17;
Matches 38; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 41
DB 215 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 255

RESULT 6
Q9BWC4 ID Q9BWC4 PRELIMINARY; PRT; 745 AA.
AC Q9BWC4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE JUNCTION PLAKOGLOBIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA, CHORIOCARCINOMA;
RA Strausberg R.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA, CHORIOCARCINOMA;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC000441; AAH00441.1; -

DR EMBL; BC011865; AAH11865.1; -
DR HSP; Q02248; 2BCT.
DR InterPro: IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 11.
DR SMART; SM00185; ARM; 12.
DR PROSITE; PS0176; ARM_REPEAT; 7.
SQ SEQUENCE 745 AA; 81726 MW; 34DF7BFB4748BCF4 CRC64;

Query Match 93.6%; Score 191; DB 4; Length 745;
Best Local Similarity 92.7%; Pred. No. 5.7e-17;
Matches 38; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 41
DB 215 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 255

RESULT 7
P70565 ID P70565 PRELIMINARY; PRT; 745 AA.
AC P70565;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PLAKOGLOBIN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=VAGINA;
RA Umekita Y., Liao S.;
RT "Molecular cloning and sequencing of the rat plakoglobin cDNA."
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=VAGINA;
RA Hiipakka R.A.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U58858; AAB06317.1; -
DR HSP; Q02248; 2BCT.
DR InterPro: IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 11.
DR SMART; SM00185; ARM; 8.
DR PROSITE; PS0176; ARM_REPEAT; 7.
SQ SEQUENCE 745 AA; 81777 MW; 9E2F52910A7ACD41 CRC64;

Query Match 93.6%; Score 191; DB 11; Length 745;
Best Local Similarity 92.7%; Pred. No. 5.7e-17;
Matches 38; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 41
DB 215 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 255

RESULT 8
Q9PVF7 ID Q9PVF7 PRELIMINARY; PRT; 729 AA.
AC Q9PVF7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CELL-ADHESION PROTEIN PLAKOGLOBIN.
GN JUP.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;

RN SEQUENCE FROM N.A.
 RP MEDLINE-99386700; PubMed=10456847;
 RA Cerda J., Reidenbach S., Pratzel S., Franke W.W.;
 RT "Caderhin-catenin complexes during zebrafish oogenesis: heterotypic
 junctions between oocytes and follicle cells.";
 RL Biol. Reprod. 61:692-704(1999).
 DR EMBL: AF099738; AAD56592.1; -
 DR HSP; Q02248; IDOW;
 DR ZFIN: ZDB-GENE-991207-22; jup.
 DR InterPro; IPR000225; Armadillo.
 DR Pfam; PF00514; Armadillo_seg; 11.
 DR SMART; SM00185; ARM; 8.
 DR PROSITE; PS50176; ARM_REPEAT; 8.
 SQ SEQUENCE 729 AA; 80033 MW; 91E00417B4FD8CEE CRC64;

 Query Match 90.2%; Score 184; DB 13; Length 729;
 Best Local Similarity 92.5%; Pred. No. 4.7e-16;
 Matches 37; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

 QY 2 REGGLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 41
 Db 205 REGGLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 244

 RESULT 9
 Q9NL44 PRELIMINARY; PRT; 769 AA.
 AC Q9NL44;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE BETA-CATENIN.
 GN CIBETA-CATENIN.
 OS Ciona intestinalis.
 OC Eukaryota; Metazoa;
 OC Cnidaria; Ciona.
 OX NCBI_TaxID=7719;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Imai K., Takada N., Satoh N., Satou Y.;
 RT "An essential role of beta-catenin in the endoderm specification of
 ascidian embryo.";
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB031543; BAA92185.1; -
 DR HSP; P35222; IG3J.
 DR InterPro; IPR000225; Armadillo.
 DR Pfam; PF00514; Armadillo_seg; 11.
 DR SMART; SM00185; ARM; 11.
 DR PROSITE; PS50176; ARM_REPEAT; 6.
 SQ SEQUENCE 769 AA; 84703 MW; F61CC489B436E1BC CRC64;

 Query Match 88.7%; Score 181; DB 5; Length 769;
 Best Local Similarity 82.9%; Pred. No. 1.3e-15;
 Matches 34; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

 QY 1 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 41
 Db 214 HKOGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 254

 RESULT 10
 O61229 PRELIMINARY; PRT; 821 AA.
 ID O61229;
 AC O61229;
 DT 01-AUG-1998 (Tremblrel. 07, Created)
 DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE BETA-CATENIN.
 OS Lytechinus variegatus (Sea urchin).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;

OC Echinozoa; Echinozoa; Echinacea; Temnopneustidae;
 OC Lytechinus.
 OX NCBI_TaxID=7654;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Miller J.R., McClay D.R.;
 RT "Changes in the pattern of adherens junction-associated beta-catenin
 accompany morphogenesis in the sea urchin embryo.";
 RL Dev. Biol. 192:310-322(1997).
 DR EMBL: U34814; AAC06340.1; -
 DR HSP; Q02248; IDOW;
 DR InterPro; IPR000225; Armadillo.
 DR Pfam; PF00514; Armadillo_seg; 12.
 DR SMART; SM00185; ARM; 12.
 DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
 DR PROSITE; PS50176; ARM_REPEAT; 8.
 SQ SEQUENCE 821 AA; 89558 MW; 71E21D562A99C5AD CRC64;

 Query Match 88.2%; Score 180; DB 5; Length 821;
 Best Local Similarity 85.4%; Pred. No. 1.8e-15;
 Matches 35; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

 QY 1 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 41
 Db 241 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 281

 RESULT 11
 Q25100 PRELIMINARY; PRT; 806 AA.
 ID Q25100;
 AC Q25100;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE BETA-CATENIN.
 OS Hydra magnipapillata (Hydra).
 OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Anthomedusae;
 OC Hydrozoa; Hydra.
 OX NCBI_TaxID=6085;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-WILDTYPE 105;
 RX MEDLINE-96257271; PubMed=8654977;
 RA Hohnmayer E., Hatta M., Fischer R., Fujisawa T., Holstein T.W.,
 RA Sugiyama T.;
 RT "Identification of a Hydra homologue of the beta-
 catenin/plakoglobin/armadillo gene family.";
 RL Gepe 172:155-159(1996).
 DR EMBL: U36781; AAC47137.1; -
 DR HSP; Q02248; 2BCT.
 DR InterPro; IPR000225; Armadillo.
 DR Pfam; PF00514; Armadillo_seg; 10.
 DR SMART; SM00185; ARM; 10.
 DR PROSITE; PS50176; ARM_REPEAT; 7.
 SQ SEQUENCE 806 AA; 90462 MW; 689E5E982CD5051A CRC64;

 Query Match 78.4%; Score 160; DB 5; Length 806;
 Best Local Similarity 73.2%; Pred. No. 8.2e-13;
 Matches 30; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

 QY 1 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 41
 Db 283 HREGLLAIFKSGGIPALVKMLGSPVDSVLFYAITTLHNL 323

 RESULT 12
 O44326 PRELIMINARY; PRT; 678 AA.
 ID O44326;
 AC O44326;

[illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 16:07:52 ; Search time 158.47 Seconds
(without alignments)
28.737 Million cell updates/sec

Title: US-09-641-104A-7
Perfect score: 203
Sequence: 1 HQBGAMVAVLAGLQKRWAL.....NKTNVKFLAITDCLQILAY 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*

- 1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	203	100.0	41	20	AA1980.DAT:*
2	192.5	94.8	42	20	AA1981.DAT:*
3	192.5	94.8	781	21	AA1982.DAT:*
4	192.5	94.8	781	21	AA1983.DAT:*
5	192.5	94.8	781	22	AA1984.DAT:*
6	192.5	94.8	781	22	AA1985.DAT:*
7	192.5	94.8	781	22	AA1986.DAT:*
8	186.5	91.9	800	22	AA1987.DAT:*
9	163.5	80.5	840	22	AA1988.DAT:*
10	163.5	80.5	840	22	AA1989.DAT:*
11	163.5	80.5	840	22	AA1990.DAT:*

12	61	30.0	41	20	AA1980.DAT:*
13	60	29.6	41	20	AA1981.DAT:*
14	57	28.1	1673	21	AA1982.DAT:*
15	56	27.6	959	22	AA1983.DAT:*
16	54	26.6	835	21	AA1984.DAT:*
17	54	26.6	845	21	AA1985.DAT:*
18	54	26.6	923	21	AA1986.DAT:*
19	53.5	26.4	682	22	AA1987.DAT:*
20	52.5	25.9	349	20	AA1988.DAT:*
21	52.5	25.9	359	22	AA1989.DAT:*
22	52.5	25.9	421	20	AA1990.DAT:*
23	52	25.6	270	19	AA1991.DAT:*
24	52	25.6	291	22	AA1992.DAT:*
25	52	25.6	669	22	AA1993.DAT:*
26	52	25.6	908	21	AA1994.DAT:*
27	51	25.1	109	20	AA1995.DAT:*
28	50	24.6	240	22	AA1996.DAT:*
29	50	24.6	262	20	AA1997.DAT:*
30	50	24.6	463	22	AA1998.DAT:*
31	50	24.6	557	22	AA1999.DAT:*
32	50	24.6	1000	21	AA2000.DAT:*
33	50	24.6	1119	22	AA2001.DAT:*
34	50	24.6	4472	17	AA2002.DAT:*
35	49	24.1	94	21	AA2003.DAT:*
36	49	24.1	101	21	AA2004.DAT:*
37	49	24.1	113	21	AA2005.DAT:*
38	48.5	23.9	305	18	AA2006.DAT:*
39	48	23.6	96	21	AA2007.DAT:*
40	48	23.6	116	21	AA2008.DAT:*
41	48	23.6	118	21	AA2009.DAT:*
42	48	23.6	140	21	AA2010.DAT:*
43	48	23.6	140	21	AA2011.DAT:*
44	48	23.6	142	21	AA2012.DAT:*
45	48	23.6	731	21	AA2013.DAT:*

ALIGNMENTS

RESULT 1

AA1980.DAT:*

ID AA1980.DAT: standard; peptide; 41 AA.

AC AA1980.DAT:*

DT 18-NOV-1999 (first entry)

XX Human beta-catenin protein armadillo repeat arm4 fragment.

DE Beta-catenin; human; armadillo repeat; treatment; human disease; LEP-1; modulator; interaction domain; transcription factor; TCF-4; oncogenic; tumor suppressor gene; conductin; E-cadherin; screening; carcinoma; APC; Wnt signalling pathway; tumor development; anti-oncogenic; melanoma; organ, regeneration; tissue regeneration; hair growth.

OS Homo sapiens.

XX DE19909251-A1.

PN 26-AUG-1999.

XX 22-FEB-1999; 99DE-1009251.

PR 21-FEB-1998; 98DE-1007390.

PA (DELB-) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.

PI Birchmeier W, Von Kries J;

XX WPI; 1999-470389/40.

DR Agents for treating human diseases, particularly cancer, modulate interaction of beta-catenin with transcription factors or tumor

Human beta-catenin
Human beta-catenin
Cosmid CHIRIM5 enco
Drosophila melanog
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Human protein sequ
Mycobacterium spec
Mycobacterium spec
Mycobacterium spec
Di-Tripeptide tran
Novel human diago
Drosophila melanog
Human ORF ORF2275
Chlamydia pneumoni
Drosophila melanog
Amino acid sequenc
Propionibacterium
Drosophila melanog
Human ORF ORF1178
Novel human diago
Virulence gene clu
Arabidopsis thalia
Arabidopsis thalia
Sulfolobus solfata
Zea mays protein f
Zea mays protein f
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Zea mays protein f
Human ORF ORF1377

suppressor gene products

Example 4: Fig. 5: 16pp; German.

This invention describes a novel agent (A) for treating human disease which is based on substances (I) that modulate (inhibit or promote) the interaction of beta-catenin with transcription factors or products of tumor-suppressor genes. The invention also describes (a) peptides (II) comprising part of the LEF-1/TCF-4 transcription factors, or their variants and mutants; (b) peptides and related molecules (III) from the armadillo domain (arm units 3-8) of beta-catenin, and mutants of the complete beta-catenin molecule, that include at least one of the specific interaction domains for LEF-1, TCF-4, APC, conductin or E-cadherin; and (c) enzyme-linked immunosorbent assay (ELISA) for screening substance libraries for compounds that modulate interaction of beta-catenin with LEF-1/TCF, APC, conductin or E-cadherin. Beta-catenin is a key compound in the Wnt signalling pathway and is involved in development of tumors. Generally its interaction with LEF-1 or TCF-4 is oncogenic but interaction with APC, conductin or E-cadherin is anti-oncogenic. (A) which inhibit interaction are particularly used to treat tumors, especially carcinoma of the colon and melanoma, but also, where they promote interaction, to stimulate regeneration of organs and tissues, specifically hair growth. AAY3230-73341 represent mutant human beta-catenin, armadillo repeat fragments described in the method of the invention.

Sequence 42 AA:

Query Match 94.8%; Score 192.5; DB 20; Length 42;
Best Local Similarity 97.6%; Pred. NO. 1.2e-22;
Matches 41; Conservative 0; Mismatches 0; Indels 1

1 HOEGA -MAVRLAGGLQKMWALLNKTNVKFLAITDCLQILAY 41

RESULT 3

AAB07290
in AAB07200 standard. protein: 781 AA.

AA
AC
AAR07290:XX
DT 01-NOV-2000 (first entry)XX
human: beta catenin
DE

β catenin; cadherin; metastasis; cadherin-associated protein; human; colorectal cancer; melanoma; antisense oligonucleotide; gene therapy.

Homo sapiens.

XX
PN USC6065500-A[illegible]XX
05 JUN - 1000.
; 09015-0344519.XX
0000-0344E10

XX
XX
XX

PA (ISIS) ISIS PHARM INC
XX
DI Bennett CE Cowsett LM

XX
PI Bennett St, Concord, MA 01734-1000
0000 430651/25

DR WP1; 2000-410651/33.
DR N-PCNB. AAA58320.

XX New antisense compounds targeting nucleic acids encoding human beta
PT catenin (HBC) useful for treating diseases associated with HBC
PT expression and as prophylaxis to prevent or delay infection.
PT Identification of tumor formation

CC or more armidillo-like repeats which participate in cadherin binding.
CC Other mutants include those lacking amino acids 555-781 or 424-781,
CC or 1-422 of beta-catenin.

XX Sequence 781 AA;
SQ

Query Match
Best Local Similarity 94.8%; Score 192.5; DB 21; Length 781;
Matches 41; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 HQEGA-MAVRLAGGLQKVVALLNKTNVFLAITTDCQLQILAY 41
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 265 hgegakmavriaggigknvllnktnkvflaittdclqilay 306

RESULT 5
AAU28118
ID AAU28118 standard; Protein: 781 AA.
XX
AC AAU28118;
XX
DT 18-DEC-2001 (first entry)
XX
XX Novel human secretory protein, Seq ID No 287.
XX
KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
KW myotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
KW gut protection; lung; liver fibrosis; immune deficiency; infection;
KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
KW fertility; analgesic; pain; antigen.
XX
OS Homo sapiens.
XX
PN WO200166689-A2.
XX
PD 13-SEP-2001.
XX
XX 05-MAR-2001; 2001WO-US04942.
XX
PR 07-MAR-2000; 2000US-0519705.
PR 19-MAY-2000; 2000US-0574454.
PR 17-JUN-2000; 2000US-0596193.
PR 14-JUL-2000; 2000US-0616847.
PR 19-SEP-2000; 2000US-0665363.
PR 20-OCT-2000; 2000US-0693267.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang Y-T, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
PI Zhao Q-A, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
XX
XX WPI: 2001-589934/66.
DR N-PSDB; AAS45018.
XX
XX Novel polypeptides and nucleic acids obtained from cDNA libraries
PT prepared from various human tissues, for diagnosis and treatment of
PT cancer, neurological, inflammatory, and autoimmune disorders -
XX
PS Example 4; SEQ ID NO 287; 107pp; English.
XX
XX The invention relates to novel isolated human secreted polypeptides (I)
XX and polynucleotides (II). (I) and (II) are useful for treating
CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
CC involved in increasing haematopoiesis, stem cell survival, bone growth
CC and remodeling. (I), (II) and modulators of (II) are useful for
CC prophylaxis or treatment of one or more cancers. (II) is also useful for
CC creating transgenic animals useful for studying the in vivo activities of

XX	781 AA;
SQ	sequence

Query Match 94.8%; Score 192.5; DB 22; Length 781;
Best Local Similarity 97.6%; Pred. No. 4.8e-21;
Matches 41; Conservative 0; Mismatches 0; Indels 1;

Db 265 hqegakmavrlaggqlqkmvallnktvkvflaittdclqllay 300

AAE06039 ID AAE06039 standard; Protein; 781 AA.
XX

DT 25-SE

DE Mouse·beta-catenin protein #2.

KW Mouse; st

KW signalling pathway: beta-c

XX Mus. musculus.

PN WO200152649-A1.
XX
PD 26-JUL-2001.

17-JAN-2001;

XX
PR
18-JAN-2000: 20XX
PA (STIR) \ INTV I.E.L.A.N.D STANFORD

XX
XX

XX

DR N-PSDB; AAD11165..

PT producing cells that retain the

PT beta-catenin in the

XX
FI
Deca calculo in ci

PS The present invention relates to a method for in vitro expansion of
CC mammalian stem or progenitor cells, which comprising increasing the
CC intracellular concentration of beta-catenin in a progenitor or stem cell

page 28-30. 33pp. English.

Sequence 41 AA:

Query Match 29.6%; Score 60; DB 20; Length 41;
Best Local Similarity 40.0%; Pred. No. 0.048;
Matches 14; Conservative 8; Mismatches 11; Indels 2; Gaps 1;

QY 1 HOEGAMAVRLAGGLQKMWALLNK--TNVKFLAITT 33
DB 1 hregllaifksggipalvkmglspvdsivfyaatt 35

RESULT 14

AAV95707
ID AAV95707 standard; Protein; 1673 AA.

AC AAV95707;

XX 25-OCT-2000 (first entry)

XX Cosmid CHRIM5 encoded protein PI4-2f.

XX Cosmid CHRIM5; nematocide; nematode; biological control agent;
KW transgenic plant; helminthiasis; PI4-2f.

XX Xenorhabdus bovienii.

XX WO200042855-A1.

XX 27-JUL-2000.

XX 24-JAN-2000; 2000WO-GB00219.

XX 22-JAN-1999; 99GB-0001499.

XX (HORT-) HORTICULTURE RES INT.

XX Morgan JAW, Jarrett P, Ellis D, Ousley MA;

XX WPI; 2000-499157/44.

XX N-PSDB; AAA50029.

XX Novel composition used to control parasitic nematodes, especially in
PT plants such as maize, cotton, soya, and rice, comprises a bacterium
PT which is a symbiont of an entomopathogenic nematode -

XX Example 6; Page 42-43; 74pp; English.

XX The present sequence is that of protein PI4-2f encoded by an open
CC reading frame identified in cosmid CHRIM5 (see AAA50029). CHRIM5 was
CC obtained by ligating Xenorhabdus bovienii strain I73 (NCIMB 40986)
CC Sau3A-digested DNA fragments into the BamHI site of the Stratagene
CC cosmid vector Supercos1, packaging into Escherichia coli XL Blue 1,
CC and screening for nematocidal activity against Caenorhabditis elegans.
CC Analysis of the DNA indicated a number of open reading frames for
CC which the corresponding protein sequences were determined (see
CC AAY95685-Y95735). Nematodes can be controlled through the use of
CC bacteria associated symbiotically with an entomopathogenic nematode.
CC Such bacteria include Xenorhabdus and Photobacterium spp. such as X.
CC bovienii strain I73. The symbiont bacteria, an engineered
CC bacterium, or a nematocidal protein obtained from such bacteria,
CC particularly PI3-1f (see AAY95706) or PI4-2f can be used to control
CC helminthiasis in a human or domesticated animal or for the control
CC of plant pathogen nematodes. Also claimed are vectors for
CC expressing nematocidal proteins in host cells, and transgenic
CC plants.

XX Sequence 1673 AA;

Query Match 28.1%; Score 57; DB 21; Length 1673;
Best Local Similarity 33.3%; Pred. No. 16;
Matches, 13; Conservative 10; Mismatches 14; Indels 2; Gaps 1;

QY P.4 GAMAVRLAGGLQKMWALLNKTNVKF--LAITTDCLQILA 40
DB 1406 ggmataaaggiaaiaastttalafgalsvtsditsivs 1444

RESULT 15

ABB58953
ID ABB58953 standard; Protein; 959 AA.

AC ABB58953;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 3651.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX Pharmaceutical;

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE.) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL03056.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

XX Disclosure; SEQ ID NO 3651; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL01840-ABL16175), expressed DNA
CC sequences (ABB57737-ABB72072),
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at:ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 959 AA;

Query Match 27.6%; Score 56; DB 22; Length 959;
Best Local Similarity 34.3%; Pred. No. 11;
Matches 12; Conservative 6; Mismatches 17; Indels 0; Gaps 0;

QY 2 OEGMAYRLAGGLQKMWALLNKTNVKFLAITDCL 36

DB 39 kngivlfrtggvghmvrllckmekimevvsil 73

Search completed: July 29, 2002, 16:07:53
Job time: 543 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 16:05:07 ; Search time 57.6 Seconds
(without alignments)
17.386 Million cell updates/sec

Title: us-09-641-104a-7
Perfect score: 203
Sequence: 1 HQEGAMVRLAGLQKRWAL.....NKTNVKFLAITDCLQLLAY 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles!.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	25.6	270	4	US-08-936-165A-487
2	50	24.6	328	2	US-08-637-759B-447
3	50	24.6	328	3	US-08-871-355A-447
4	50	24.6	328	4	US-09-201-945-447
5	48.5	23.9	305	2	US-08-602-359A-42
6	48	23.6	620	4	US-08-982-785A-10
7	48	23.6	686	4	US-08-982-785A-8
8	48	23.6	756	4	US-08-982-785A-2
9	47	23.2	423	4	US-08-939-366-6
10	47	23.2	426	6	5221737-2
11	46	22.7	318	1	US-08-190-802A-33
12	46	22.7	318	4	US-08-477-346-33
13	46	22.7	318	4	US-08-473-089-33
14	46	22.7	820	1	US-08-291-896-2
15	46	22.7	820	2	US-08-485-278-2
16	46	22.7	844	3	US-09-029-267-20
17	45	22.2	43	4	US-08-981-739-104
18	45	22.2	134	4	US-08-981-739-81
19	45	22.2	142	4	US-08-931-858E-111
20	45	22.2	142	4	US-08-981-739-111
21	45	22.2	185	4	US-08-981-739-133
22	44.5	21.9	347	4	US-09-094-557-3
23	44	21.7	791	4	US-08-861-745B-1
24	44	21.7	1015	3	US-08-463-210-9
25	44	21.7	1015	4	US-09-124-900-3
26	44	21.7	1016	2	US-07-743-357-2
27	44	21.7	1016	2	US-07-743-357-4

Sequence 85, Appl
Sequence 784, App
Sequence 25, Appl
Sequence 54, Appl
Sequence 54, Appl
Patent No. 5204232
Sequence 28, Appl
Sequence 28, Appl
Sequence 136, Appl
Sequence 19, Appl
Sequence 28, Appl
Sequence 40, Appl
Sequence 40, Appl
Sequence 40, Appl
Sequence 1, Appl
Sequence 5, Appl
Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-936-165A-487
; Sequence 487, Application US/08936165A
; Patent No. 6348582
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Burnham, Martin
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Lonetto, Michael
; APPLICANT: Nicholas, Richard
; APPLICANT: Pratt, Julie
; APPLICANT: Reichard, Richard
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
; TITLE OF INVENTION: Polypeptides and Their Uses
; NUMBER OF SEQUENCES: 534
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,165A
; FILING DATE: 24-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/027,032
; FILING DATE: 24-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50549
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 487:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 270 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

Tue Jul 30 08:15:39 2002

us-09-641-104a-7.ra

TOPOLOGY: linear
MOLECULE TYPE: Protein
US-08-936-165A-487

Query Match 25.6%; Score 52; DB 4; Length 270;
Best Local Similarity 32.6%; Pred. No. 2.8;
Matches 14; Conservative 8; Mismatches 15; Indels 1;

QY 5 AMAVRLAGGLQKMWALLNKN-----VKFLAITTDCLOILAY 41
DB 214 ASQAINGTLVKLIEPLGOTNVFIFLGVAIVITICISILTF 256

RESULT 2
US-08-637-759B-447
Sequence 447, Application US/08637759B
Patent No. 5876931
GENERAL INFORMATION:
APPLICANT: David William Holden
TITLE OF INVENTION: Identification of Genes
NUMBER OF SEQUENCES: 501
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/637,759B
FILING DATE: 03-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/02875
FILING DATE: 11-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: RPMS 101
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 447:
SEQUENCE CHARACTERISTICS:
LENGTH: 328 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-637-759B-447

Query Match 24.6%; Score 50; DB 2; Length 328;
Best Local Similarity 32.4%; Pred. No. 7.3;
Matches 12; Conservative 6; Mismatches 19; Indels 0;

QY 4 GAMAVRLAGGLQKMWALLNKNVKFLAITTDCLOILLA 40
DB 201 GFFAIQLPGGIYARVLLTDTNTMKFDELVDIETLLA 237

RESULT 3
US-08-871-355A-447
Sequence 447, Application US/08871355A

PATENT NO. 6015669
GENERAL INFORMATION:
APPLICANT: David William Holden
TITLE OF INVENTION: Identification of Genes
NUMBER OF SEQUENCES: 501
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/871,355A
FILING DATE: 09-JUN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/02875
FILING DATE: 11-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: RPMS 101 CON
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 447:
SEQUENCE CHARACTERISTICS:
LENGTH: 328 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-871-355A-447

Query Match 24.6%; Score 50; DB 3; Length 328;
Best Local Similarity 32.4%; Pred. No. 7.3;
Matches 12; Conservative 6; Mismatches 19; Indels 0;

QY 4 GAMAVRLAGGLQKMWALLNKNVKFLAITTDCLOILLA 40
DB 201 GFFAIQLPGGIYARVLLTDTNTMKFDELVDIETLLA 237

RESULT 4
US-09-201-945-447
Sequence 447, Application US/09201945
Patent No. 6342215
GENERAL INFORMATION:
APPLICANT: David William Holden
TITLE OF INVENTION: Identification of Genes
NUMBER OF SEQUENCES: 501
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

Query: Match      23.6%; Score 48; DB 4; Length 620;
Best Local Similarity 45.0%; Pred. No. 32;
Matches: 9; Conservative: 6; Mismatches 5; Indels
QY 1 HQECAMAVRLAGGLQKMWAL 20
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

Db 540 NMENAKALRDAGGIEKLVI 559

RESULT 7

US-08-982-785A-8
Sequence 8, Application US/08982785A
Patent No. 6258929
GENERAL INFORMATION:
APPLICANT: Kosik, Kenneth S.
APPLICANT: Zhou, Jianhua
TITLE OF INVENTION: ALARM RELATED PEPTIDES AND
NUCLEIC ACIDS AND DIAGNOSIS USING THEM
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/982,785A
FILING DATE: 02-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/031,556
FILING DATE: 02-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 05311/018001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 686 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-982-785A-8

Query Match 23.6%; Score 48; DB 4; Length 686;
Best Local Similarity 45.0%; Pred. No. 37;
Matches 9; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 1 HOEGAMAVRLAGGLOKQWAL 20

Db 537 NMENAKALRDAGGIEKLVI 556

RESULT 8

US-08-982-785A-2
Sequence 2, Application US/08982785A
Patent No. 6258929
GENERAL INFORMATION:
APPLICANT: Kosik, Kenneth S.
APPLICANT: Zhou, Jianhua
TITLE OF INVENTION: ALARM RELATED PEPTIDES AND
NUCLEIC ACIDS AND DIAGNOSIS USING THEM
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/939,366
FILING DATE: 29-SEP-1997
CLASSIFICATION: .435
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: OHU-02749
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 6:

ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/982,785A
FILING DATE: 02-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/031,556
FILING DATE: 02-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 05311/018001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 756 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-982-785A-2

Query Match 23.6%; Score 48; DB 4; Length 756;
Best Local Similarity 45.0%; Pred. No. 41;
Matches 9; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 1 HOEGAMAVRLAGGLOKQWAL 20

Db 541 NMENAKALRDAGGIEKLVI 560

RESULT 9

US-08-939-366-6
Sequence 6, Application US/08939366
Patent No. 6355415
GENERAL INFORMATION:
APPLICANT: Wagner, Thomas E.
APPLICANT: Xie, Yuesheng
TITLE OF INVENTION: Compositions and Methods for the Use of
Ribozymses to Determine Gene Function
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medien & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/939,366
FILING DATE: 29-SEP-1997
CLASSIFICATION: .435
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: OHU-02749
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 6:

TELEPHONE: (415) 324-

10

22.78; Score 46; DB 4; Length 318;

Best Local Similarity 35.1%; Pred. No. 29; Mismatches 11; Indels 10; Gaps 1;

QY 1 HQEGAMAVR-----LAGGLQKVMALLNKTNK 27
 DB 149 HTEWVSCVRFSPMTNPIIVSGGWDKMKVKNVNLNCK 185

RESULT 13
 US-08-473-089-33
 Sequence 33, Application US/08473089
 Patent No. 6342368
 GENERAL INFORMATION:
 APPLICANT: Mochly-Rosen, Daria
 APPLICANT: Ron, Dorit
 TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
 TITLE OF INVENTION: Thereof
 NUMBER OF SEQUENCES: 265
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Morrison & Foerster
 STREET: 2000 Pennsylvania Avenue, NW
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20006-1812
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/473,089
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: MURASHIGE, KATE H.
 REGISTRATION NUMBER: 29,959
 REFERENCE/DOCKET NUMBER: 2550-0025.22
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 887-1500
 TELEFAX: (202) 887-0763
 INFORMATION FOR SEQ ID NO: 33:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 318 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: GBLP -CHLAMIDOMONAS HOMOLOG, Fig. 16

Query Match 22.7%; Score 46; DB 4; Length 318;
 Best Local Similarity 35.1%; Pred. No. 29; Mismatches 11; Indels 10; Gaps 1;

QY 1 HQEGAMAVR-----LAGGLQKVMALLNKTNK 27
 DB 149 HTEWVSCVRFSPMTNPIIVSGGWDKMKVKNVNLNCK 185

RESULT 14
 US-08-291-896-2
 Sequence 2, Application US/08291896
 Patent No. 5480800
 GENERAL INFORMATION:
 APPLICANT:
 TITLE OF INVENTION: DNA fragment carrying the gene encoding the
 enzyme for fragmenting N-acetylheparosan and the adjacent
 sequences permitting its expression, recombinant enzyme
 TITLE OF INVENTION: intended for fragmenting N-acetylheparosan and its use.

NUMBER OF SEQUENCES: 7
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/291,896
 FILING DATE: 17-AUG-1994
 CLASSIFICATION: 435
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 820 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-291-896-2

Query Match 22.7%; Score 46; DB 1; Length 820;
 Best Local Similarity 32.0%; Pred. No. 94; Mismatches 6; Indels 11; Gaps 0;

QY 3 EGAMAVRIAGGLQKVMALLNKTNK 27
 DB 594 EGDVSLAVGGGISSQLKLFNSDNTK 618

RESULT 15
 US-08-485-278-2
 Sequence 2, Application US/08485278
 Patent No. 5820857
 GENERAL INFORMATION:
 APPLICANT: Legoux, Richard
 APPLICANT: Lelong, Philippe
 APPLICANT: Salome, Marc Louis Victor
 TITLE OF INVENTION: DNA fragment carrying the gene encoding the enzyme for
 fragmenting N-acetylheparosan and the adjacent sequences pe
 TITLE OF INVENTION: fragmenting N-acetylheparosan and the adjacent sequences pe
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Jacobson, Price, Holman & Stern
 STREET: 400 Seventh St. N.W.
 CITY: Washington, D.C.
 COUNTRY: U.S.A.
 ZIP: 20004
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/485,278
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: FR 93/10050
 FILING DATE: 17-AUG-1993
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 820 amino acids
 TYPE: amino acid

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-485-278-2

```

Query Match 22.7%; Score 46; DB 2; Length 820;
Best Local Similarity 32.0%; Pred. No. 94;
Matches 8; Conservative 6; Mismatches 11; Indels

Qy 3 EGAMAVRLAGGLQKMWALLNKTNVK 27
 || :: : || : : | | | |
 Db 594 EGDVSLAVGGGIISSQLKLFNSDNTK 618

Search completed: July 29, 2002, 16:05:08
Job time: 478 sec

THIS PAGE BLANK (USPTO)

Query Match	94.8%;	Score 192.5;	DB 2;	Length 781;
Best Local Similarity	97.6%;	Pred. No. 4.1e-18;		
Best Overall Similarity	97.6%;	N. Mismatches 0;		
Conservative	94.1%			
Indels	1;			
Gaps	1;			

3

Query Match	94.8%	Score 192.5;	DB 2;	Length 781;
Best Local Similarity	97.6%	Pred. NO. 4.1e-18;		
		0 misassemblies		
		0 indels		
		0 gaps		

FCIT. T. A.

Query Match	84.0%;	Score 170.5;	DB 2;	Length 820;
Best Local Similarity	85.7%;	Pred. No. 4.3e-15;		
Matches 16:	Conservative	3;	Mismatches 2;	Indels 1;
				Gaps 1;

RESULT

Science 257, 1142-1144, 1992
A:Title: Plakoglobin and beta-catenin: distinct but closely related.
A:Reference number: S35091; MUID:92376536
A:Accession: S35092
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-621 <BUT>
A:Cross-references: EMBL:M90365
C:Keywords: cytoskeleton

Query Match 79.6%; Score 161.5; DB 2; Length 621;
Best Local Similarity 83.3%; Pred. No. 5.4e-14;
Matches 35; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 1 HOEGA-MAVRLAGGLQKQKVVALLNKTNVKFLAITTDCLQILAY 41
DB 132 YQEGAKMAVRLADGLQKQKVVPLLNKNNPKFLAITTDCLQLAY 173

RESULT 8

S35093
Plakoglobin - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jul-2000
C:Accession: S35093; S24636
R:Fouquet, B.; Zimbelmann, R.; Franke, W.W.
Differentiation 51, 187-194, 1992
A:Title: Identification of plakoglobin in oocytes and early embryos of Xenopus laevis.
A:Reference number: S35093; MUID:93093332
A:Accession: S35093
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-738 <FOU>
A:Cross-references: EMBL:M95593; NID:9214656; PIDN:AAA49931.1; PID:g214657
R:Demarais, A.A.
submitted to the EMBL Data Library, June 1992
A:Reference number: S24636
A:Accession: S24636
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 133-184, 'V', 186-225, 'T', 227-292 <DEM>
A:Cross-references: EMBL:X67078; NID:965252; PID:g65253
C:Keywords: cytoskeleton

Query Match 79.6%; Score 161.5; DB 2; Length 738;
Best Local Similarity 83.3%; Pred. No. 6.5e-14;
Matches 35; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 1 HOEGA-MAVRLAGGLQKQKVVALLNKTNVKFLAITTDCLQILAY 41
DB 252 YQEGAKMAVRLADGLQKQKVVPLLNKNNPKFLAITTDCLQLAY 293

RESULT 9

JC4835
beta-catenin - Hydra magnipapillata
C:Species: Hydra magnipapillata
C:Date: 15-Aug-1996 #sequence_revision 15-Oct-1996 #text_change 21-Jul-2000
C:Accession: JC4835
R:Hohmayer, E.; Hatta, M.; Fischer, R.; Fujisawa, T.; Holstein, T.W.; Sugiyama, T.
Gene 172, 155-159, 1996
A:Title: Identification of a Hydra homologue of the beta-catenin/plakoglobin/armadillo
A:Reference number: JC4835; MUID:96257271
A:Accession: JC4835
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-806 <HOB>
A:Cross-references: GB:U36781; NID:g1407600; PIDN:AA47137.1; PID:g1407601
C:Comment: This protein plays roles in cadherin-mediated cell adhesion and in signal tra
C:Genetics:
A:Gene: betaCtn

Query Match 75.6%; Score 153.5; DB 2; Length 806;
Best Local Similarity 81.0%; Pred. No. 8.8e-13;
Matches 34; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 1 HOEGA-MAVRLAGGLQKQKVVALLNKTNVKFLAITTDCLQILAY 41
DB 324 HOEGAKMAVRLALGLQKQKVVLLQRPKVKFLAIVTDCQLILAY 365

RESULT 10

A32905
plakoglobin, desmosomal - human
C:Species: Homo sapiens (man)
C:Date: 22-Nov-1989 #sequence_revision 22-Nov-1989 #text_change 29-Aug-1997
C:Accession: A32905
R:Franke, W.W.; Goldschmidt, M.D.; Zimbelmann, R.; Mueller, H.M.; Schiller, D.L.; Cow
Proc. Natl. Acad. Sci. U.S.A. 86, 4027-4031, 1989
A:Title: Molecular cloning and amino acid sequence of human plakoglobin, the common j
A:Reference number: A32905; MUID:89264555
A:Accession: A32905
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-744 <FRA>
A:Cross-references: GB:M23410
C:Genetics:
A:Gene: GDB:JUP
A:Cross-references: GDB:I26565; OMIM:173325
A:Map position: 7pter-7qter
C:Keywords: cytoskeleton

Query Match 62.1%; Score 126; DB 2; Length 744;
Best Local Similarity 65.9%; Pred. No. 4.6e-09;
Matches 27; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

QY 1 HOEGAMAVRLAGGLQKQKVVALLNKTNVKFLAITTDCLQILAY 41
DB 256 YQEGAKMACAGRRQKQKVVPLLNKNNPKFLAITTDCLQLAY 296

RESULT 11

T11032
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 - redhead mitochondrion
C:Species: mitochondrion Aythya americana (redhead)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 07-Dec-1999
C:Accession: T11032
R:Mindell, D.P.; Sorenson, M.D.; Dimcheff, D.E.
Proc. Natl. Acad. Sci. U.S.A. 95, 10693-10697, 1998
A:Title: Multiple independent origins of mitochondrial gene order in birds.
A:Reference number: Z17242
A:Accession: T11032
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-607 <JOH>
A:Cross-references: EMBL:AF090337; NID:g4887659; PID:g4887670; PIDN:AA32262.1
C:Genetics:
A:Gene: mitochondrion
A:Superfamily: NADH dehydrogenase (ubiquinone) chain 5
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation

Query Match 27.3%; Score 55.5; DB 2; Length 607;
Best Local Similarity 40.0%; Pred. No. 15;
Matches 14; Conservative 2; Mismatches 8; Indels 11; Gaps 1;

QY 4 GAMAVRLAGGLQKQKVVALLNKTNVKFLAITTDCLQI 38
DB 353 SPQDIRKMGGLQKQK-----LPVTTSLTI 376

RESULT 12

A72865
AcOrf-120 protein - Autographa californica nuclear polyhedrosis virus
C:Species: Autographa californica nuclear polyhedrosis virus, ACMPPV
A:Note: dsDNA virus
C:Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 12-Nov-1999
C:Accession: A72865
R:Ayres, M.D.; Howard, S.C.; Kuzio, J.; Lopez-Ferber, M.; Possee, R.D.
Virology 202, 586-605, 1994
A:Title: The complete DNA sequence of Autographa californica nuclear polyhedrosis virus.
A:Reference number: A72850; MUID:94303173
A:Accession: A72865
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-82 <A>
A:Cross-references: GB:L22858; NID:g510708; PIDN:AAA66750.1; PID:g559189
C:Genetics:
A:Gene: AcOrf-120

Query Match 26.4%; Score 53.5; DB 2; Length 82;
Best Local Similarity 40.0%; Pred. No. 3.6;
Matches 12: Conservative 6; Mismatches 7; Indels 5; Gaps 1;

QY 14 LQKMVALLNKTNVKFLAI-----TTDCLQI 38
| : | ||| : | : ||| :
Db 33 LPNLELLRRKNTNIKEVAVMLNKKTTTECLQL 62

RESULT 13

T41857
ACMPV orf120 - Bombyx mori nuclear polyhedrosis virus (isolate T3)
C:Species: Bombyx mori nuclear polyhedrosis virus, BmSNPV
A:Variety: Isolate T3
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000
C:Accession: T41857
R:Gomi, S.; Majima, K.; Maeda, S.
J. Gen. Virol. 80, 1323-1337, 1999
A:Title: Sequence analysis of the genome of Bombyx mori nucleopolyhedrovirus.
A:Reference number: Z2020; MUID:99281911
A:Accession: T41857
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-82 <KAM>
A:Cross-references: EMBL:L33180; PIDN:AAC63786.1
A:Experimental source: isolate T3
C:Genetics:
A:Note: Orf 98

Query Match	26.4%	Score 53.5;	DB 2;	Length 82;
Best Local Similarity	40.0%	Pred. NO. 3.6;		
Matches 12: Conservative	6;	Mismatches 7;	Indels 5;	Gaps 1;

QY 14 LQKMWALLNKTNVKFLAI-----TTDCLQI 38
| : | | | | : | : | | | | :
Db 33 LDNLKLLRRKTNIKEVAVMLNKKTTTECLQI 62

RESULT 14

AH2264
 two-component hybrid sensor and regulator alr3671 [imported] - *Anabaena* sp. (strain PCC 7120)
 C:Species: *Anabaena* sp.
 A:Note: *Anabaena* sp. (strain PCC 7120) is a synonym of *Nostoc* sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
 C:Accession: AH2264
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena* sp. Strain PCC 7120
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AH2264
 A:Status: preliminary

A;Molecule: type: DNA
A;Residues: 1-407 <KUR>

A;Residues: 1-407 <KUR>
A;CrossReferences: 1GB:BAO00019; PIDN:BAE75370.1; PID:gl7132804; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: [glr3671](#)

Query Match

Best Local Similarity: 22.5%; Pred. NO. 19;
Matches: 9; Conservative 17; Mismatches 13; Indels 1; Gaps 1;

2 0EGAMAVRLAGGLOKMAVLLNKTNVKFLAITTDCLQILAY 41

1827 OCCAIG-NLSPOMOEVIAMARSNINLLSMVNTLLEVYRE 220

RESULT 15

G81146
50S ribosomal protein L25 NWB0876 [imported] - *Neisseria meningitidis* (strain MC58 se
C:Species: *Neisseria meningitidis*
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: G81146
R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.
ri, H.; Qiu, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Massignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;
A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
A:Reference number: A81000: MUID:20175755

A; Status: preliminary

A; Molecule type: DNA

A; Residues: 1-190 <TET>

A; Cross-references: GB:AE002098; NID:g/226112; PIDN:AAF4128/.1; PID:g/z
A-Experimental source: serogroup B, strain MC58

A; Experiment
C: Genetics:

A:Gene: NMB0876

Query Match 26.1%; Score 53; DB 2; Length 190;

Best Local Similarity 34.8%; pred. No. 10;

Matches	16;	Conservative	10;	Mismatches	8;	Indels	12;	Gaps	3;
---------	-----	--------------	-----	------------	----	--------	-----	------	----

QV 3 EGAMAVRLAGGLQKMVALLNKTNVKFLAITT-----DCLQILA 40

113 ENSOAVKLOGG---RVSLN-TSVEVVALPANIPAFELDLCAEVA 154

search completed: July 29, 2002, 16:09:11

Job time: 721 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 16:09:50 ; Search time 34.24 Seconds
(without alignments)
46.364 Million cell updates/sec

Title: US-09-641-104a-7
Perfect score: 203
Sequence: 1 HQGAMAVRLAGLQKMWAL.....NKTNVKFLAITTDCLOLAY 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	192.5	94.8	781	1 CTNB_HUMAN	P35222 homo sapien
2	192.5	94.8	781	1 CTNB_MOUSE	Q02248 mus musculus
3	192.5	94.8	781	1 CTNB_RAT	Q9WU82 rattus norv
4	192.5	94.8	781	1 CTNB_XENLA	P26233 xenopus lae
5	170.5	84.0	820	1 CTNB_TRIGR	P35223 tripneustes
6	167.5	82.5	818	1 CTNB_URECA	P35224 urechis cau
7	163.5	80.5	813	1 ARM_MUSDO	Q02453 musca domes
8	163.5	80.5	843	1 ARM_DROME	P18824 drosophila
9	161.5	79.6	621	1 PLAK_MOUSE	Q02257 mus musculus
10	161.5	79.6	738	1 PLAK_XENLA	P30998 xenopus lae
11	126	62.1	743	1 PLAK_HUMAN	P14923 homo sapien
12	53.5	26.4	82	1 Y120_NPVAC	P41673 autographa
13	52.5	25.9	349	1 YP63_WCTU	Q50735 mycobacteri
14	52	25.6	945	1 AB83_HUMAN	Q60423 homo sapien
15	51.5	25.4	319	1 ORC6_ARATH	Q92VH3 arabidopsis
16	50.5	24.9	775	1 ICPO_HSV11	P08393 herpes simp
17	50	24.6	256	1 6PCL_CHLTR	O84189 chlamydia t
18	50	24.6	322	1 SSQK_SALTY	P74860 salmonella
19	50	24.6	933	1 AB82_HUMAN	P98198 homo sapien
20	49.5	24.4	690	1 EP42_HUMAN	P16452 homo sapien
21	48.5	23.9	605	1 NUSM_STRCA	Q35813 struthio ca
22	48	23.6	113	1 RK20_CYPAP	P14809 cyanophora
23	48	23.6	729	1 AB82_HUMAN	Q9at12 homo sapien
24	48	23.6	781	1 PBPB_HAETN	P45345 h penicilli
25	48	23.6	1148	1 AB82_MOUSE	P98200 mus musculus
26	47.5	23.4	686	1 EP42_BOVIN	Q46510 bos taurus
27	47.5	23.4	2869	1 TRP1_PLAVB	Q00798 plasmodium
28	47	23.2	253	1 TRP1_LACLA	Q01997 lactococcus
29	47	23.2	409	1 KP8S_SCHPO	P41831 schizosacch
30	47	23.2	423	1 BRAC_BRARE	Q07998 brachydanio
31	47	23.2	423	1 SYH_SALTY	O52765 salmonella
32	47	23.2	426	1 GABT_ECOLI	P22256 escherichia
33	47	23.2	492	1 NU4M_CHOCR	P48915 chondrus cr

34. 47 23.2 1581 1 ACC8_CRICR Q09427 cricetus cr
35. 47 23.2 3066 1 POLG_SEMVG Q90069 s genome po
36. 46 22.7 178 1 PGSA_AQUAE O67908 aquifex aeo
37. 46 22.7 233 1 DEOD_TREPA O83716 treponema p
38. 46 22.7 318 1 GBLP_CHLRE P25387 chlamydomon
39. 46 22.7 335 1 YJ89_PHRHO O57713 pyrococcus
40. 46 22.7 340 1 MOAA_STACA Q924m6 staphylococ
41. 46 22.7 350 1 GP40_BFSP1 O48396 bacterioph
42. 46 22.7 365 1 GRAB_BACSU P07869 bacillus su
43. 46 22.7 371 1 DHAS_BUCAI P57523 buchnera ap
44. 46 22.7 421 1 DHOM_HELPY P56429 helicobacte
45. 46 22.7 460 1 HYN_ERWHE Q47860 erwina her

ALIGNMENTS

RESULT 1
CTNB_HUMAN STANDARD; PRT; 781 AA.
ID CTNB_HUMAN
AC P35222:
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Beta-catenin.
GN CTNNB1 OR CTNNB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=95105247; PubMed=7806582;
RA Huelksen J., Birchmeier W., Behrens J.;
RT "E-cadherin and APC compete for the interaction with beta-catenin and
the cytoskeleton".
RL J. Cell Biol. 127:2061-2069(1994).
RN [2]
RP REVIEW.
RX MEDLINE=20145417; PubMed=10679188;
RA Kikuchi A.;
RT "Regulation of beta-catenin signaling in the Wnt pathway".
RL Biochem. Biophys. Res. Commun. 268:243-248(2000).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 133-664.
RX MEDLINE=20578806; PubMed=11136974;
RA Graham T.A., Weaver C., Mao F., Kimelman D., Xu W.;
RT "Crystal structure of a beta-catenin/Tcf complex".
RL Cell 103:885-896(2000).
CC -!- FUNCTION: INVOLVED IN THE REGULATION OF CELL ADHESION AND IN
SIGNAL TRANSDUCTION THROUGH THE WNT PATHWAY.
CC -!- SUBUNIT: THE CYTOSOLIC FORM BOUND TO AXIN ALLOWS THE STIMULATION
OF ITS DEGRADATION. THE CYTOSOLIC FORM ALSO INTERACTS WITH APC AND
E-CADHERIN. THE NUCLEAR FORM BINDS TCF/LEF-1 AND MAY ALSO BIND
PONTIN2 AND DUPLIN. CAN HETERODIMERIZE WITH ALPHA-CATENIN AND
ALSO FORM A TRIMERIC COMPLEX WITH CADHERIN. MEDIATES THE ANCHORAGE
OF THE CADHERIN COMPLEX TO THE CORTICAL ACTIN CYTOSKELETON.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC WHEN IT IS UNSTABILIZED (HIGH
LEVEL OF PHOSPHORYLATION). TRANSLOCATES TO THE NUCLEUS WHEN IT IS
STABILIZED (LOW LEVEL OF PHOSPHORYLATION).
CC -!- PTM: PROBABLY PHOSPHORYLATED BY GSK-3B. WNT-DEPENDENT ACTIVATION
OF DVL ANTAGONIZES THE ACTION OF GSK-3B. PHOSPHORYLATION OF BETA-
CATENIN IS REDUCED AND THE PROTEIN IS NO LONGER DEGRADED,
RESULTING IN ITS ACCUMULATION IN CYTOPLASM.
CC -!- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -!- SIMILARITY: CONTAINS 12 ARM REPEATS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial

entitles requires a license agreement (see <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC EMBL; X87838; CAA61107.1; -;
 CC EMBL; Z19054; CAA79497.1; -;
 CC PIR; S31988; S31988.
 CC PDB; 1G3J; 17-JAN-01.
 CC TRANSFAC; T02872; -;
 CC MIM; 116806; -;
 CC InterPro; IPR000225; Armadillo.
 CC Pfam; PF00514; Armadillo_seg; 12.
 CC SMART; SM00185; ARM; 11.
 CC PROSITE; PS50176; ARM_REPEAT; 9.
 KW Cell adhesion; Cytoskeleton; Structural protein; Nuclear protein;
 KW Repeat; 3D-structure.
 FT REPEAT 151 191 ARM 1.
 FT REPEAT 193 234 ARM 2.
 FT REPEAT 235 276 ARM 3.
 FT REPEAT 277 318 ARM 4.
 FT REPEAT 319 360 ARM 5.
 FT REPEAT 361 389 ARM 6.
 FT REPEAT 400 441 ARM 7.
 FT REPEAT 442 484 ARM 8.
 FT REPEAT 489 530 ARM 9.
 FT REPEAT 531 571 ARM 10.
 FT REPEAT 594 636 ARM 11.
 FT REPEAT 637 666 ARM 12.
 SQ SEQUENCE 781 AA; 85496 MW; CB78F165A3EEF86E CRC64;
 Query Match 94.8%; Score 192.5; DB 1; Length 781;
 Best Local Similarity 97.6%; Pred. No. 6.5e-19;
 Matches 41; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1 HQEGA-MAVRLAGLGLOKMKVALLNKTNVFLAITTDCLOILAY 41
 Db 265 HQEAKMAVRLAGLGLOKMKVALLNKTNVFLAITTDCLOILAY 306
 RESULT 2
 CTNNB_MOUSE STANDARD; PRT; 781 AA.
 AC Q02248;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Beta-catenin.
 GN CTNNB1 OR CATNB.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92376536; PubMed=1509266;
 RA Butz S., Stappert J., Weissig H., Kemler R.;
 RT "Plakoglobin and beta-catenin: distinct but closely related";
 RL Science 257:1142-1144(1992).
 RN [2]
 RP REVIEW.
 RX MEDLINE=20145417; PubMed=10679188;
 RA Kikuchi A.;
 RT "Regulation of beta-catenin signaling in the Wnt pathway";
 RL Biochem. Biophys. Res. Commun. 268:243-248(2000).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 150-665.
 RX MEDLINE=97442350; PubMed=9298899;
 RA Huber A.H., Nelson W.J., Weiss W.I.;
 RT "Three-dimensional structure of the armadillo repeat region of beta-catenin";
 RL Cell 90:871-882(1997).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 118-149 IN COMPLEX WITH

RP CTNNB1.
 RX MEDLINE=20337986; PubMed=10882138;
 RA Pokutta S., Weis W.I.;
 RT "Structure of the dimerization and beta-catenin-binding region of alpha-catenin";
 RL Mol. Cell 5:533-543(2000).
 CC -1- FUNCTION: INVOLVED IN THE REGULATION OF CELL ADHESION AND IN SIGNAL TRANSDUCTION THROUGH THE WNT PATHWAY.
 CC -1- SUBUNIT: THE CYTOSOLIC FORM BOUND TO AXIN ALLOWS THE STIMULATION OF ITS DEGRADATION. THE CYTOSOLIC FORM ALSO INTERACTS WITH APC AND E-CADHERIN. THE NUCLEAR FORM BINDS TCF/LEF-1 AND MAY ALSO BIND PONTIN52 AND DUPLIN. CAN HETERODIMERIZE WITH ALPHA-CATENIN AND ALSO FORM A TRIMERIC COMPLEX WITH CADHERIN. MEDIATES THE ANCHORAGE OF THE CADHERIN COMPLEX TO THE CORTICAL ACTIN CYTOSKELETON.
 CC -1- OF THE CELLULAR LOCATION: CYTOPLASMIC WHEN IT IS UNSTABILIZED (HIGH LEVEL OF PHOSPHORYLATION). TRANSLOCATES TO THE NUCLEUS WHEN IT IS STABILIZED (LOW LEVEL OF PHOSPHORYLATION).
 CC -1- PTM: PROBABLY PHOSPHORYLATED BY GSK-3B. WNT-DEPENDENT ACTIVATION OF DVL ANTAGONIZES THE ACTION OF GSK-3B. PHOSPHORYLATION OF BETA-CATENIN IS REDUCED AND THE PROTEIN IS NO LONGER DEGRADED, RESULTING IN ITS ACCUMULATION IN CYTOPLASM.
 CC -1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 12 ARM REPEATS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).
 CC EMBL; M90364; AAA37280.1; -;
 DR PIR; S35091; S35091.
 DR PDB; 2BCT; 15-OCT-97.
 DR PDB; 3BCT; 19-NOV-97.
 DR PDB; 1DOW; 12-JUL-00.
 DR TRANSFAC; T02984; -;
 DR MGD; MGI:88276; Catnb.
 DR InterPro; IPR000225; Armadillo.
 DR Pfam; PF00514; Armadillo_seg; 12.
 DR SMART; SM00185; ARM; 11.
 DR PROSITE; PS50176; ARM_REPEAT; 9.
 DR Cell adhesion; Cytoskeleton; Structural protein; Nuclear protein;
 KW Repeat; 3D-structure.
 FT REPEAT 151 191 ARM 1.
 FT REPEAT 193 234 ARM 2.
 FT REPEAT 235 276 ARM 3.
 FT REPEAT 277 318 ARM 4.
 FT REPEAT 319 360 ARM 5.
 FT REPEAT 361 389 ARM 6.
 FT REPEAT 400 441 ARM 7.
 FT REPEAT 442 484 ARM 8.
 FT REPEAT 489 530 ARM 9.
 FT REPEAT 531 571 ARM 10.
 FT REPEAT 594 636 ARM 11.
 FT REPEAT 637 666 ARM 12.
 SQ SEQUENCE 781 AA; 85470 MW; D708F170A3FBED6E CRC64;

Query Match 94.8%; Score 192.5; DB 1; Length 781;
 Best Local Similarity 97.6%; Pred. No. 6.5e-19;
 Matches 41; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1 HQEGA-MAVRLAGLGLOKMKVALLNKTNVFLAITTDCLOILAY 41
 Db 265 HQEAKMAVRLAGLGLOKMKVALLNKTNVFLAITTDCLOILAY 306
 RESULT 3
 CTNNB_MOUSE STANDARD; PRT; 781 AA.
 ID CTNNB_MOUSE
 AC Q9WU82;

DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Beta-catenin.
 GN CTNNB1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Testis;
 RX MEDLINE=99428593; PubMed=10497305;
 RA Chung S.W., Lee W.M., Cheng C.Y.;
 RT "Study on the formation of specialized inter-Sertoli cell junctions in
 vitro.";
 RL J. Cell. Physiol. 181:258-272(1999).
 CC -1- FUNCTION: INVOLVED IN THE REGULATION OF CELL ADHESION AND IN
 CC SIGNAL TRANSDUCTION THROUGH THE WNT PATHWAY (BY SIMILARITY).
 CC -1- SUBUNIT: THE CYTOSOLIC FORM BOUND TO AXIN ALLOWS THE STIMULATION
 CC OF ITS DEGRADATION. THE CYTOSOLIC FORM ALSO INTERACTS WITH APC AND
 CC E-CADHERIN. THE NUCLEAR FORM BINDS TCF/LEF-1 AND MAY ALSO BIND
 CC PONTIN52 AND DUPLIN. CAN HETERODIMERIZE WITH ALPHA-CATENIN AND
 CC ALSO FORM A TRIMERIC COMPLEX WITH CADHERIN. MEDIATES THE ANCHORAGE
 CC OF THE CADHERIN COMPLEX TO THE CORTICAL ACTIN CYTOSKELETON (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC WHEN IT IS UNSTABILIZED (HIGH
 CC LEVEL OF PHOSPHORYLATION). TRANSLOCATES TO THE NUCLEUS WHEN IT IS
 CC STABILIZED (LOW LEVEL OF PHOSPHORYLATION) (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE TESTIS.
 CC -1- DEVELOPMENTAL STAGE: HIGHLY EXPRESSED AT E30-60 DAY DPC IN THE
 CC TESTIS. REDUCED EXPRESSION AT E90 DAY DPC.
 CC -1- PTM: PROBABLY PHOSPHORYLATED BY GSK-3B. WNT-DEPENDENT ACTIVATION
 CC OF DVL ANTAGONIZES THE ACTION OF GSK-3B. PHOSPHORYLATION OF BETA-
 CC CATENIN IS REDUCED AND THE PROTEIN IS NO LONGER DEGRADED,
 CC RESULTING IN ITS ACCUMULATION IN CYTOPLASM (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 12 ARM REPEATS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF121265; AAD28504.1; -
 DR HSSP; Q02248; 2BCT.
 DR InterPro; IPR000225; Armadillo.
 DR Pfam; PF00514; Armadillo_seg; 12.
 DR SMART; SM00185; ARM; 11.
 DR PROSITE; PS50176; ARM_REPEAT; 9.
 KW Cell adhesion; Cytoskeleton; Structural protein; Nuclear protein;
 Repeat.
 FT REPEAT 151 191 ARM 1.
 FT REPEAT 193 234 ARM 2.
 FT REPEAT 235 276 ARM 3.
 FT REPEAT 277 318 ARM 4.
 FT REPEAT 319 360 ARM 5.
 FT REPEAT 361 389 ARM 6.
 FT REPEAT 400 441 ARM 7.
 FT REPEAT 442 484 ARM 8.
 FT REPEAT 489 530 ARM 9.
 FT REPEAT 531 571 ARM 10.
 FT REPEAT 594 636 ARM 11.
 FT REPEAT 637 666 ARM 12.
 SQ SEQUENCE 781 AA; 85454 MW; 9C29186B5DD54B87 CRC64;
 Query Match 94.8%; Score 192.5; DB 1; Length 781;
 Best Local Similarity 97.6%; Pred. No. 6.5e-19;
 Matches 41; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 HOEGA-MAVRLAGGQKMWALLKNTNKFLLATTTDCLQILAY 41
 DB 265 HOEGA-MAVRLAGGQKMWALLKNTNKFLLATTTDCLQILAY 306
 RESULT 4.
 CTNB_XENLA STANDARD; PRT; 781 AA.
 AC P26233; 1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22; Last sequence update)
 DT 01-MAR-2002 (Rel. 41; Last annotation update)
 DE Beta-catenin.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RC TISSUE=Brain.
 RX MEDLINE=92073903; PubMed=1962194;
 RA McCrea P.D., Turk C.W., Gumbiner B.M.;
 RT "A homolog of the armadillo protein in Drosophila (plakoglobin)
 associated with E-cadherin.";
 RL Science 254:1359-1361(1991).
 CC -1- FUNCTION: BINDS TO THE CYTOPLASMIC DOMAIN OF THE CELL-CELL
 CC ADHESION MOLECULE E-CADHERIN, AND PERHAPS TO OTHER (MEMBRANE)
 CC PROTEINS. THE ASSOCIATION OF CATENINS TO CADHERINS PRODUCES A
 CC COMPLEX WHICH IS LINKED TO THE ACTIN FILAMENT NETWORK, AND WHICH
 CC SEEMS TO BE OF PRIMARY IMPORTANCE FOR CADHERINS CELL-ADHESION
 CC PROPERTIES. ESSENTIAL COMPONENT OF THE VERTEBRATE WNT-SIGNALING
 CC PATHWAY AND IS REQUIRED FOR FORMATION OF THE EMBRYONIC AXIS.
 CC -1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 10 ARM REPEATS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M77013; AAA49670.1; -
 DR PIR; S35099; S35099.
 DR HSSP; Q02248; 2BCT.
 DR TRANSFAC; T03026; -
 DR InterPro; IPR000225; Armadillo.
 DR Pfam; PF00514; Armadillo_seg; 12.
 DR SMART; SM00185; ARM; 11.
 DR PROSITE; PS50176; ARM_REPEAT; 9.
 KW Cell adhesion; Cytoskeleton; Structural protein; Repeat.
 FT REPEAT 141 180 ARM 1.
 FT REPEAT 225 264 ARM 2.
 FT REPEAT 267 306 ARM 3.
 FT REPEAT 351 390 ARM 4.
 FT REPEAT 391 429 ARM 5.
 FT REPEAT 432 473 ARM 6.
 FT REPEAT 479 519 ARM 7.
 FT REPEAT 521 562 ARM 8.
 FT REPEAT 584 623 ARM 9.
 FT REPEAT 625 664 ARM 10.
 SQ SEQUENCE 781 AA; 85449 MW; 3ECD27232239F799 CRC64;
 Query Match 94.8%; Score 192.5; DB 1; Length 781;
 Best Local Similarity 97.6%; Pred. No. 6.5e-19;
 Matches 41; Conservative 0; Mismatches 0; Indels 1; Gaps 1;


```

RP SEQUENCE FROM N.A.
RX MEDLINE-93247062; PubMed-8483160;
RA Pelfer M.A., Wieschaus E.;
RT "The product of the Drosophila melanogaster segment polarity gene
RT armadillo is highly conserved in sequence and expression in the
RT housefly Musca domestica.";
RL J. Mol. Evol. 36:224-233(1993).
CC -1- FUNCTION: SEGMENT POLARITY PROTEIN. MUTATION IN ARM PRIMARILY
CC AFFECT THE POSTERIOR PART OF THE SEGMENT AND LEAD TO THE
CC PRODUCTION OF ANTERIOR STRUCTURES WITHIN THIS REGION. MAY BIND
CC A CADHERIN AND PARTICIPATE IN THE TRANSMISSION OF DEVELOPMENTAL
CC INFORMATION.
CC -1- SUBCELLULAR LOCATION: INNER SURFACE OF CELL MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -1- SIMILARITY: CONTAINS 10 ARM REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L04874; AAA29292.1; --
CC DR HSSP; Q02248; IDOW.
CC DR InterPro; IPR000225; Armadillo.
CC DR Pfam; PF00514; Armadillo_seg; 12.
CC DR SMART; SM00185; ARM; 11.
CC DR PROSITE; PS50176; ARM_REPEAT; 9.
CC KW Developmental protein; Segmentation polarity protein; Repeat;
CC Cell adhesion; Cytoskeleton; Structural protein.
CC FT DOMAIN 1 148 ASP/GLU-RICH (ACIDIC).
CC FT REPEAT 139 178 ARM 1.
CC FT REPEAT 180 221 ARM 2.
CC FT REPEAT 223 262 ARM 3.
CC FT REPEAT 265 304 ARM 4.
CC FT REPEAT 349 387 ARM 5.
CC FT REPEAT 388 427 ARM 6.
CC FT REPEAT 430 471 ARM 7.
CC FT REPEAT 477 517 ARM 8.
CC FT REPEAT 587 626 ARM 9.
CC FT REPEAT 628 667 ARM 10.
CC SQ SEQUENCE 813 AA; 88237 MW; 73E61F59BDBFA580 CRC64;

Query Match 80.5%; Score 163.5; DB 1; Length 813;
Best Local Similarity 81.08; Pred. No. 7e-15;
Matches 34; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY 1 HQEGA-MAVRLAGGLQKRWALLNKNVFLAITTDCILQILAY 41
||: ||||| ||||| ||: ||||| ||||| |||||
Db 263 HDGSKAVRLAGGLQKRWVLLQNNVFLAIVTDCILQILAY 304
||: ||||| ||||| ||: ||||| ||||| |||||

RESULT 8
ARM_DROME STANDARD; PRT; 843 AA.
AC P18824; Q09546;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Armadillo segment polarity protein.
OS ARM OR EG:86E4.6 OR CG11579.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN
RP SEQUENCE FROM N.A. (CYTOPLASMIC ISOFORM).
RC STRAIN=OREGON-R;
RX MEDLINE-89211895; PubMed-2707602;

```

```

RA Riggleman B., Wieschaus E., Schedl P.;
RT "Molecular analysis of the armadillo locus: uniformly distributed
RT transcripts and a protein with novel internal repeats are associated
RT with a Drosophila segment polarity gene.";
RL Genes Dev. 3:96-113(1989).
[2]
RP SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, AND ALTERNATIVE
RP SPLICING.
RC TISSUE=Head;
RX MEDLINE-98298928; PubMed-9635189;
RA Loureiro J., Pelfer M.;
RT "Roles of Armadillo, a Drosophila catenin, during central nervous
RT system development.";
RL Curr. Biol. 8:622-632(1998).
[3]
RP SEQUENCE FROM N.A. (CYTOPLASMIC ISOFORM).
RC STRAIN=BERKELEY;
RX MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Escler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
[4]
RP SEQUENCE FROM N.A. (CYTOPLASMIC ISOFORM).
RC STRAIN=OREGON-R;
RX MEDLINE-20196011; PubMed-10731137;
RA Beros P.V., Gatt M.K.; Ashburner M., Murphy L., Harris D.,
RA Barrell B.G., Ferraz C., Vidal S., Brun C., Demailles J., Cadieu E.,
RA Dreano S., Gloux S., Lelaure V., Mottier S., Galibert F., Borkova D.,
RA Mirana B., Kafatos F.C., Louis C., Siden-Kiamos I., Bolshakov S.,
RA Papagiannakis G., Spanos L., Cox S., Madueno E., de Pablos B.,
RA Modolell J., Peter A., Schoettler P., Werner M., Mourikoti F.,
RA Beinert N., Dowe G., Schaefer U., Jaekle H., Bucheton A.,
RA Callister D.M., Campbell L.A., Darlamitsou A., Henderson N.S.,
RA McWilliam P.J., Salles C., Tait E.A., Valenti P., Saunders R.D.C.,
RA Glover D.M.;
RT "From sequence to chromosome: the tip of the X chromosome of D.
RT melanogaster.";
RL Science 287:2220-2222(2000).

```

[5]
RN PHOSPHORYLATION.
RX MEDLINE=95113174; PubMed=7529201;
RT Peifer M., Pai L.-M., Casey M.;
RT "Phosphorylation of the Drosophila adherens junction protein
RT Armadillo: roles for wingless signal and zeste-white 3 kinase."
RL Dev. Biol. 166:543-556(1994).
CC -1- FUNCTION: NEURAL ISOFORM MAY ASSOCIATE WITH CADN AND PARTICIPATE
CC IN THE TRANSMISSION OF DEVELOPMENTAL INFORMATION. CAN ASSOCIATE
CC WITH ALPHA-CATENIN. CYTOPLASMIC ISOFORM ACCUMULATES THROUGH WG
CC SIGNALING. ARM FUNCTION IN WG SIGNAL TRANSDUCTION IS REQUIRED
CC EARLY IN DEVELOPMENT FOR DETERMINATION OF NEUROBLAST FATE. ARM AND
CC ABL PROTEINS FUNCTION COOPERATIVELY AT ADHERENS JUNCTIONS IN BOTH
CC THE CNS AND EPIDERMIS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC AND MEMBRANE-ASSOCIATED WITH THE
CC INNER SURFACE OF CELL MEMBRANE.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; CYTOPLASMIC (SHOWN HERE) AND
CC NEURAL; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: CYTOPLASMIC ISOFORM IS DOMINANT BEFORE GERM
CC BAND RETRACTION, AFTER RETRACTION AND DURING LARVAL STAGES, IT IS
CC FOUND IN HIGH LEVELS IN SPECIFIC CELLS ALONG THE CNS MIDLINE.
CC NEURAL ISOFORM IS FIRST SEEN AFTER GERM BAND RETRACTION IN THE
CC AXON TRACTS OF THE CNS, ALSO PRESENT IN AXONS DURING LARVAL STAGES
CC AND ACCUMULATES IN THE MOTOR NEURONS OF THE SEGMENTAL AND
CC INTERSEGMENTAL NERVES AS THEY EXIT THE CNS. BOTH ISOFORMS
CC ACCUMULATE IN THE PNS.
CC -1- DEVELOPMENTAL STAGE: PRESENT AT ALL STAGES, BUT REACHED THE
CC HIGHEST LEVELS DURING EARLY TO MID-EMBRYOGENESIS.
CC -1- PTM: PHOSPHORYLATED ON SER, THR AND TYR RESIDUES. LEVEL OF
CC PHOSPHORYLATION VARIES BOTH DURING EMBRYONIC DEVELOPMENT AND FROM
CC EMBRYONIC TISSUE TO TISSUE. SGG IS REQUIRED FOR PHOSPHORYLATION
CC AND WG SIGNAL NEGATIVELY REGULATES ARM PHOSPHORYLATION.
CC HYPHOPHOSPHORYLATED FORM OF ARM INCREASES IN STEADY-STATE LEVELS.
CC -1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -1- SIMILARITY: CONTAINS 12.5 ARM REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X54468; CAA38350.1; .
CC EMBL; AF001213; AAB58731.1; .
CC EMBL; AF003422; AAF45688.1; ALT_INIT.
CC EMBL; AL021106; CAA15946.1; .
CC EMBL; AL021086; CAA15946.1; JOINED.
CC EMBL; AL021086; CAA15935.1; .
CC EMBL; AL021106; CAA15935.1; JOINED.
CC PIR; A31861; A31861.
CC HSSP; Q02248; IDOW.
CC TRANSFAC; T02977; .
CC FlyBase; FBgn0000117; arm.
CC InterPro; IPR000225; Armadillo.
CC Pfam; PF00514; Armadillo_seg; 12.
CC SMART; SM00185; ARM; 11.
CC PROSITE; PS50176; ARM_REPEAT; 9.
CC Developmental protein; Segmentation polarity protein; Repeat;
KW Cell adhesion; Cytoskeleton; Structural protein; Phosphorylation;
KW Alternative splicing.
FT DOMAIN 1 158 ASP/GLU-RICH (ACIDIC).
FT REPEAT 159 200
FT REPEAT 201 242
FT REPEAT 243 284
FT REPEAT 285 326
FT REPEAT 327 368
FT REPEAT 369 410
FT REPEAT 411 449
FT REPEAT 450 496
FT REPEAT 497 538
FT REPEAT 539 584
FT REPEAT 585 608
FT REPEAT 609 647
FT REPEAT 648 689
FT DOMAIN 690 843
FT VARSPTIC 718 843
FT
FT
SQ SEQUENCE 843 AA; 91152 MW; 40DAD6FB83163049 CRC64;
Query Match 80.58; Score 163.5; DB 1; Length 843;
Best Local Similarity 81.08; Pred. No. 7.3e-15;
Matches 34; Conservative 3; Mismatches 4; Indels 1; Gaps 1;
QY HOEGA-MAVRLAGGLQKVMALLNKTNVFLAITTDCLQLAY 41
DB 273 HDGSKMAVRLAGGLQKVMVTLQRNNVFLAIVTDCQLQLAY 314
RESULT 9
ID PLAK_MOUSE STANDARD; PRT; 621 AA.
AC Q02257;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Junction plakoglobin (Desmoplakin III) (Fragment).
GN JUP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92376536; PubMed=1509266;
RA Butz S., Stappert J., Weissig H., Kemler R.;
RT "Plakoglobin and beta-catenin: distinct but closely related."
RL Science 257:1142-1144(1992).
RP REVISIONS TO 294 AND 296.
RA Butz S.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: COMMON JUNCTIONAL PLAQUE PROTEIN. THE MEMBRANE-
CC ASSOCIATED PLAQUES ARE ARCHITECTURAL ELEMENTS IN AN IMPORTANT
CC STRATEGIC POSITION TO INFLUENCE THE ARRANGEMENT AND FUNCTION
CC OF BOTH THE CYTOSKELETON AND THE CELLS WITHIN THE TISSUE. THE
CC PRESENCE OF PLAQUEGLOBIN IN BOTH THE DESMOSOMES AND IN THE
CC INTERMEDIATE JUNCTIONS SUGGESTS THAT IT PLAYS A CENTRAL ROLE IN
CC THE STRUCTURE AND FUNCTION OF SUBMEMBRANOUS PLAQUES.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC IN A SOLUBLE AND MEMBRANE-
CC ASSOCIATED FORM.
CC -1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -1- SIMILARITY: CONTAINS AT LEAST 9 ARM REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M90365; AAB02885.1; .
CC PIR; S35092; S35092.
CC HSSP; Q02248; IDOW.
CC MGD; MGI:96650; Jup.
CC InterPro; IPR000225; Armadillo.
CC Pfam; PF00514; Armadillo_seg; 11.
CC SMART; SM00185; ARM; 8.
CC PROSITE; PS50176; ARM_REPEAT; 9.
KW Cell adhesion; Cytoskeleton; Structural protein; Repeat.
KW

```

EMBL: X67078; CAA47463.1; -
PIR: S24636; S24636.
PIR: S35093; S35093.
HSSP: Q02248; 2BCT.
InterPro: IPR000225; Armadillo.
Pfam: PF00514; Armadillo_seg; 11.
SMART: SM00185; ARM; 10.
PROSITE: PS50176; ARM_REPEAT; 9.
KW Cell adhesion; Cytoskeleton; Structural protein; Repeat.
FT REPEAT 128 167 ARM 1.
FT REPEAT 212 251 ARM 2.
FT REPEAT 254 293 ARM 3.
FT REPEAT 338 377 ARM 4.
FT REPEAT 379 416 ARM 5.
FT REPEAT 419 460 ARM 6.
FT REPEAT 466 506 ARM 7.
FT REPEAT 508 547 ARM 8.
FT REPEAT 570 609 ARM 9.
FT REPEAT 611 650 ARM 10.
FT CONFLICT 185 185 I -> V (IN REF. 2).
FT CONFLICT 226 226 A -> T (IN REF. 2).
FT SEQUENCE 738..AA; 81711 MW; 569DBE69D08BBC58 CRC64;
Query Watch 79.6%; Score 161.5; DB 1; Length 738;
Best Local Similarity 83.3%; Pred. No. 1.2e-14;
Matches 35; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
QY 1 HQGCA-MAYRLAGLGKMWALLNKTIVKFLAITDCLQILAY 41
|||||
DB 252 YQEGAKMAVRLADGLQKMWPLLNNKPKFLAITDCLQILAY 293
|||||

RESULT: 11
PLAK_HUMAN:
ID ID PLAK_HUMAN STANDARD; PRT; 743 AA.
AC AC P14923;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Junction plakoglobin (Desmoplakin III).
GN Jup,OR dp3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RS SEQUENCE FROM N.A.
RP MEDLINE=8926455; PubMed=2726765;
RA Franke W.W., Goldschmidt M.D., Zimbelmann R., Mueller H.M.,
RA Schiller D.L., Cowin P.;
RT "Molecular cloning and amino acid sequence of human plakoglobin, the
RT common junctional plaque protein."
RL Proc.:Natl. Acad. Sci. U.S.A. 86:4027-4031(1989).
CC -1- FUNCTION: COMMON JUNCTIONAL PLAQUE PROTEIN. THE MEMBRANE-
CC ASSOCIATED PLAQUES ARE ARCHITECTURAL ELEMENTS IN AN IMPORTANT
CC STRATEGIC POSITION TO INFLUENCE THE ARRANGEMENT AND FUNCTION
CC OF BOTH THE CYTOSKELETON AND THE CELLS WITHIN THE TISSUE. THE
CC PRESENCE OF PLAKOGLOBIN IN BOTH THE DESMOSES AND IN THE
CC INTERMEDIATE JUNCTIONS SUGGESTS THAT IT PLAYS A CENTRAL ROLE IN
CC THE STRUCTURE AND FUNCTION OF SUBMEMBRANOUS PLAQUES.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC IN A SOLUBLE AND MEMBRANE-
CC ASSOCIATED FORM.
CC -1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -1- SIMILARITY: CONTAINS 8 ARM REPEATS.
CC -1- SIMILARITY: CONTAINS 8 ARM REPEATS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/

```

or send an email to license@lsb-sib.ch).

CC EMBL: M23410; AAA64895.1; --
 CC FIR: A32905; A32905.
 DR HSSP: Q02248; 2BCT.
 DR MIM: 173325; --
 DR InterPro: IPR000225; Armadillo.
 DR Pfam: PF00514; Armadillo_seg; 10.
 DR SMART: SM00185; ARM; 8.
 DR PROSITE: PS50176; ARM_REPEAT; 8.
 KW Cell adhesion; Cytoskeleton; Structural protein; Repeat.
 FT INIT_MET 0 0 PROBABLE.
 FT REPEAT 141 179 ARM 1.
 FT REPEAT 183 224 ARM 2.
 FT REPEAT 225 252 ARM 3.
 FT REPEAT 308 351 ARM 4.
 FT REPEAT 389 430 ARM 5.
 FT REPEAT 431 473 ARM 6.
 FT REPEAT 478 521 ARM 7.
 FT REPEAT 582 624 ARM 8.
 FT CONFLICT 96 99 GOL A -> DSSL (IN REF. 1; AAA64895).
 FT CONFLICT 139 139 V -> A (IN REF. 1; AAA64895).
 FT SEQUENCE 743 AA; 81498 MW; 472741F400D388FD CRC64;

Query Match 62.1%; Score 126; DB 1; Length 743;

Best Local Similarity 65.9%; Pred. No. 1e-09; Mismatches 12; Indels 0; Gaps 0;

QY 1 HQSGAMVRIAGGLQKMKVALLNTNKKFLAITTDCQLLAY 41
 DB 255 YQEGAKMACAGRAQKMKVPLLNKPNKFLAITTDCQLLAY 295

RESULT 12

ID Y120_NPVAC STANDARD; PRT; 82 AA.
 AC P41673;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Hypothetical 9.5 kDa protein in HE65-PK2 intergenic region.
 OS Autographa californica nuclear polyhedrosis virus (AcMNPV).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID=46015;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C6;
 RX MEDLINE=94303173; PubMed=8030224;
 RA Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;
 RT "The complete DNA sequence of Autographa californica nuclear polyhedrosis virus";
 RL Virology 202:586-605(1994).
 CC -1- SIMILARITY: TO CORRESPONDING ORF IN OPMNPV.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).

CC EMBL: L22858; AAA66750.1; --
 CC Hypothetical protein.
 KW SEQUENCE 82 AA; 9532 MW; 3303CBDBCB9175F2 CRC64;
 SQ
 Query Match 26.4%; Score 53.5; DB 1; Length 82;
 Best Local Similarity 40.0%; Pred. No. 1.1;
 Mismatches 12; Conservative 6; Mismatches 7; Indels 5; Gaps 1;

QY 14 LQKVVALLNKNVKKFLAI-----TTDCQLQI 38
 DB 33 LDNLELLRRKTKNIKEVAVMLNKKTTTECLQL 62
 RESULT 13
 YP63_MYCTU STANDARD; PRT; 349 AA.
 ID YP63_MYCTU
 AC Q50735;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 36.0 kDa protein RV2563 precursor.
 GN RV2563 OR MT2639 OR MTCY9C4.05C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 Hornsby T., Jagels K., Krogh A., McLean L.A., Moule S., Murphy L.,
 Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: SOME, TO M_JANNASCHII MJ1507.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).

CC EMBL: Z77250; CAB01052.1; --
 CC EMEL; AE007098; AAK46952.1; --
 CC TIGR; MT2639; --
 CC TubercuList; RV2563;
 CC InterPro: IPR003838; DUF214.
 CC Pfam: PF02687; DUF214; 1
 KW Hypothetical protein; Transmembrane; Signal; Complete proteome.
 FT SIGNAL 1 28 POTENTIAL.
 FT CHAIN 29 349 HYPOTHETICAL PROTEIN RV2563.
 FT TRANSMEM 55 75 POTENTIAL.
 FT TRANSMEM 158 178 POTENTIAL.
 FT TRANSMEM 230 250 POTENTIAL.
 FT TRANSMEM 261 281 POTENTIAL.
 FT TRANSMEM 284 304 POTENTIAL.
 FT TRANSMEM 308 328 POTENTIAL.
 SQ SEQUENCE 349 AA; 36025 MW; 694A971AC96A53F1 CRC64;

CC or send an email to license@lsb-sib.ch).

CC -----
DR EMBL: AC005508; AAD14490.1; -
KW Hypothetical protein; DNA replication; Nuclear protein; DNA-binding.
SQ SEQUENCE 319 AA; 36181 MW; 8573C6DD26206DE8 CRC64;

Query Match 25.4%; Score 51.5; DB 1; Length 319;
Best Local Similarity 27.3%; Pred. No. 8.5;
Matches 12; Conservative 11; Mismatches 10; Indels 11; Gaps 1;

QY 7 AVRLAG-----GLQKMWALLNKTNYKFLAITTDCLOIL 39
||:|:|
|:|:|
Db 68 AVKLSGMEKAYSRFNSIQNVIGIKIKLVRELAVQFCVVRVI 111

Search completed: July 29, 2002, 16:09:52
Job time: 757 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 16:12:02 ; Search time 124.4 seconds
(without alignments)
57.016 Million cell updates/sec

Title: US-09-641-104A-7
Perfect score: 203
Sequence: 1 HOEGAMAVRLAGGLQKVAL.....NKTNVKFLAITTCQLILAY 41

Scoring table: BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	192.5	94.8	780	13	Q90424
2	192.5	94.8	781	11	Q9D335
3	192.5	94.8	781	13	O42486
4	166.5	82.0	821	5	O61229
5	161.5	79.6	745	4	Q15151
6	161.5	79.6	745	4	Q9BWC4
7	161.5	79.6	745	11	P70565
8	160.5	79.1	769	5	Q9NL44
9	157	77.3	117	6	P79321
10	156.5	77.1	729	13	Q9PVF7
11	156.5	77.1	773	5	O76152
12	153.5	75.6	806	5	Q25100
13	57	28.1	1680	2	Q9EVR7
14	56	27.6	959	5	Q9VAH2
15	55.5	27.3	607	8	Q9XL01
16	54	26.6	911	10	Q9LW87

17	53.5	26.4	82	12	O92476
18	53.5	26.4	505	4	Q9H0C0
19	53	26.1	190	16	Q9JZW3
20	53	26.1	379	16	Q9HZQ7
21	53	26.1	571	10	O04250
22	52.5	25.9	359	5	Q9V7Q2
23	52.5	25.9	445	12	Q910F9
24	52	25.6	190	16	Q9JUX7
25	52	25.6	347	2	Q9FC35
26	52	25.6	669	5	Q9VM21
27	51.5	25.4	521	10	Q9S264
28	51.5	25.4	605	8	Q9XL35
29	51.5	25.4	606	8	Q9Z228
30	51	25.1	422	2	Q9F5A8
31	51	25.1	526	10	O22266
32	51	25.1	913	10	Q9CAA0
33	51	25.1	920	10	Q9SSD5
34	50.5	24.9	228	12	O69078
35	50.5	24.9	370	10	Q9SSB7
36	50.5	24.9	399	16	Q9AB94
37	50	24.6	240	5	Q9VSH1
38	50	24.6	284	16	Q9S666
39	50	24.6	336	10	Q943X2
40	50	24.6	408	16	P74265
41	50	24.6	555	5	Q961R9
42	50	24.6	557	5	Q9VEX8
43	50	24.6	765	4	Q96143
44	50	24.6	1273	10	Q9LX99
45	49.5	24.4	1616	2	Q9KKB3

ALIGNMENTS

RESULT 1.

Q90424 ID Q90424 PRELIMINARY; PRT; 780 AA.
AC Q90424;
DT 01-NOV-1996 (TREMREL. 01, Created)
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)
DT 01-DEC-2001 (TREMREL. 19, Last annotation update)
DE B-CATENIN.
GN CTNNB1;
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96122902; PubMed=8562427;
RA Kelly G.M., Erezylmaz D.F., Moon R.T.;
RT "Induction of a secondary embryonic axis in zebrafish occurs following the overexpression of beta-catenin.";
RL Mech. Dev. 53:261-273(1995).
DR EMBL:U01081; AAC59732.1;
DR HSPF:P35222; I33J.
DR ZFIN:ZDB-GENE-980526-362; cttnb.
DR InterPro: IPR000225; Armadillo.
DR Pfam: PF00514; Armadillo_seg; 12.
DR SMART: SM00185; ARM; 11.
DR PROSITE: PS50176; ARM_REPEAT; 8.
SQ SEQUENCE 780 AA; 85542 MW; D7A1FB80F94066DC CRC64;

Query Match 94.8%; Score 192.5; DB 13; Length 780;
Best Local Similarity 97.6%; Pred. No. 2.1e-18;
Matches 41; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 HOEGAMAVRLAGGLQKVALNKTNVFLAITTCQLILAY 41
DB 264 HOEGAMAVRLAGGLQKVALNKTNVFLAITTCQLILAY 305

```

RESULT 2
Q9D335 ID Q9D335 PRELIMINARY; PRT; 781 AA.
AC Q9D335;
DT 01-JUN-2001 (TREMREL. 17, Created)
DT 01-JUN-2001 (TREMREL. 17, Last sequence update)
DT 01-DEC-2001 (TREMREL. 19, Last annotation update)
DE ADULT MALE COLON CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,
DE CLONE: 9030417H18, FULL INSERT SEQUENCE.
DE CATNB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=COLON;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK018515; BAB31250.1; -
DR HSSP; P35222; IG3J.
DR MGD; MGI:88276; Catnb.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 12.
DR SMART; SM00185; ARM; 12.
DR PROSITE; PS0176; ARM_REPEAT; 7.
SQ SEQUENCE 781 AA; 85546 MW; 937538C3B5CD75D1 CRC64;

Query Match 94.8%; Score 192.5; DB 11; Length 781;
Best Local Similarity 97.6%; Pred. No. 2.1e-18;
Matches 41; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 HQEGA-MAVRLAGLQKMKVALLNKTNVFKLAITTCQLQILAY 41
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 265 HQEGAKMAVRLAGLQKMKVALLNKTNVFKLAITTCQLQILAY 306

RESULT 3
Q42486 ID Q42486 PRELIMINARY; PRT; 781 AA.
AC Q42486;
DT 01-JAN-1998 (TREMREL. 05, Created)
DT 01-JAN-1998 (TREMREL. 05, Last sequence update)
DT 01-DEC-2001 (TREMREL. 19, Last annotation update)
DE BETA CATENIN.
DE CHECAT.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=WHITE LEGHORN; TISSUE=DORSAL SKIN;
RX MEDLINE=97464068; PubMed=9322759;
RA Li J., Chuong C.M., Widelitz R.B.;
RT "Isolation and characterization of chicken beta-catenin.";
RL Gene 196:201-207(1997).
DR EMBL; U082964; AAB80856.1; -
DR HSSP; P35222; IG3J.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 12.
DR SMART; SM00185; ARM; 11.
DR PROSITE; PS0176; ARM_REPEAT; 8.
SQ SEQUENCE 781 AA; 85438 MW; 6D205D9A4DBAC562 CRC64;

Query Match 94.8%; Score 192.5; DB 13; Length 781;
Best Local Similarity 97.6%; Pred. No. 2.1e-18;
Matches 41; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 HQEGA-MAVRLAGLQKMKVALLNKTNVFKLAITTCQLQILAY 41
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 265 HQEGAKMAVRLAGLQKMKVALLNKTNVFKLAITTCQLQILAY 306

RESULT 4
O61229 ID O61229 PRELIMINARY; PRT; 821 AA.
AC O61229;
DT 01-AUG-1998 (TREMREL. 07, Created)
DT 01-AUG-1998 (TREMREL. 07, Last sequence update)
DT 01-JUN-2001 (TREMREL. 17, Last annotation update)
DE BETA-CATENIN.
OS Lytechinus variegatus (Sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Temnopneustida; Toxopneustidae;
OC Lytechinus.
OX NCBI_TaxID=7654;
RN [1]
SEQUENCE FROM N.A.
RC MEDLINE=98104237; PubMed=9441670;
RA Miller J.R., McClay D.R.;
RT "Changes in the pattern of adherens junction-associated beta-catenin
RT accompany morphogenesis in the sea urchin embryo.";
RL Dev. Biol. 192:310-322(1997).
DR EMBL; U34814; AAC06340.1; -
DR HSSP; Q02248; IDOW.
DR InterPro; IPR002086; Aldehyde_dehydr.
DR Pfam; PF00514; Armadillo_seg; 12.
DR SMART; SM00185; ARM; 12.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
DR PROSITE; PS0176; ARM_REPEAT; 8.
SQ SEQUENCE 821 AA; 89558 MW; 71E21D562A99C5AD CRC64;

Query Match 82.0%; Score 166.5; DB 5; Length 821;
Best Local Similarity 83.3%; Pred. No. 9.9e-15;
Matches 35; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

QY 1 HQEGA-MAVRLAGLQKMKVALLNKTNVFKLAITTCQLQILAY 41
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 282 HQEGSKMAVRLAGLQKMAALLSRNPNPKFLAITTCQLQILAY 323

RESULT 5
Q15151 ID Q15151 PRELIMINARY; PRT; 745 AA.
AC Q15151; O15093;
DT 01-NOV-1996 (TREMREL. 01, Created)
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)
DT 01-JUN-2001 (TREMREL. 17, Last annotation update)
DE PLAKOGLOBIN.
OX NCBI_TaxID=9031;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=WHITE LEGHORN; TISSUE=DORSAL SKIN;
RX MEDLINE=97464068; PubMed=9322759;
RA Li J., Chuong C.M., Widelitz R.B.;
RT "Isolation and characterization of chicken beta-catenin.";
RL Gene 196:201-207(1997).
DR EMBL; U082964; AAB80856.1; -
DR HSSP; P35222; IG3J.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 12.
DR SMART; SM00185; ARM; 11.
DR PROSITE; PS0176; ARM_REPEAT; 8.
SQ SEQUENCE 781 AA; 85438 MW; 6D205D9A4DBAC562 CRC64;

Query Match 94.8%; Score 192.5; DB 11; Length 781;
Best Local Similarity 97.6%; Pred. No. 2.1e-18;
Matches 41; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 HQEGA-MAVRLAGLQKMKVALLNKTNVFKLAITTCQLQILAY 41
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 265 HQEGAKMAVRLAGLQKMKVALLNKTNVFKLAITTCQLQILAY 306

```

```

Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL: BC000441; RAH00441.1; -
DR EMBL: BC011865; AHA11865.1; -
DR HSSP: Q02248; 2BCT
DR InterPro: IPR000225; Armadillo
DR Pfam: PF00514; Armadillo_seg; 11.
DR SMART: SM00185; ARM; 12.
DR PROSITE: PS0176; ARM_REPEAT; 7.
SQ SEQUENCE 745-AA; 81726 NW; 34DF7BF478BCF4 CRC64;

Query Match          79.6%; Score 161.5; DB 4; Length 745;
Best Local Similarity 83.3%; Pred. No. 4.5e-14;
Matches   35; Conservative    2; Mismatches    4; Indels    1; Gaps

QY      1 HQEGA-MAYVLAGLGLOKVVALLNKTNVKFLAITTDCLOIILAY 41
       :||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     256 YQEGAKMAYRLADGLQKWVPLLNKNPKFLAITTDCLOQLLAY 297

RESULT      7
P70565      PRELIMINARY; PRT; 745 AA.
AC P70565;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PLAKOGLOBIN
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=SPRAGUE-DAWLEY; TISSUE=VAGINA;
RA Umekita Y., Liao S.;
RT "Molecular cloning and sequencing of the rat plakoglobin cDNA.";
RL Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.

P70565      PRELIMINARY; TISSUE=VAGINA;
AC P70565;
DT 01-FEB-1996 (MAX-1996) to the EMBL/GenBank/DDBJ databases.
DL EMBL: U58858; AAB06317.1; -.
DR HSSP: Q02248; 2BCT
DR InterPro: IPR000225; Armadillo
DR Pfam: PF00514; Armadillo_seg; 11.
DR SMART: SM00185; ARM; 8.
DR PROSITE: PS0176; ARM_REPEAT; 7.
SQ SEQUENCE 745 AA; 81777 NW; 9E2F52910A7ACD41 CRC64;

Query Match          79.6%; Score 161.5; DB 11; Length 745;
Best Local Similarity 83.3%; Pred. No. 4.5e-14;
Matches   35; Conservative    2; Mismatches    4; Indels    1; Gaps

QY      1 HQEGA-MAYVLAGLGLOKVVALLNKTNVKFLAITTDCLOIILAY 41
       :||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     256 YQEGAKMAYRLADGLQKWVPLLNKNPKFLAITTDCLOQLLAY 297

RESULT      8
Q9NL44      PRELIMINARY; PRT; 769 AA.
ID Q9NL44;
AC Q9NL44;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE BETA-CATENIN..
GN CIBETA-CATENIN.
OS Ciona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia
OC Clonidae; Ciona.
```

```

OX NCBI_TaxID=7719;
RN [1]
RP SEQUENCE FROM N.A.
RA Imai K., Takada N., Satoh N., Satou Y.;
RT "An essential role of beta-catenin in the endoderm specification of
RT ascidian embryo.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB031543; BAA92185.1; -.
DR HSSP; P35222; IG3J.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 11.
DR SMART; SM00185; ARM; 11.
DR PROSITE; PS50176; ARM_REPEAT; 6.
SQ SEQUENCE 769 AA; 84703 MW; F61CC489B436E1BC CRC64;

Query Match 79.1%; Score 160.5; DB 5; Length 769;
Best Local Similarity 83.3%; Pred. No. 6.4e-14;
Matches 35; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 1 HOEGAM-AVRLAGGLQKMWVLLNKNVNFKLAITTDCLQILAY 41
Db 255 HOEGAKEAVRLAGGLQKMWVLLSRDNRVNFKLAITTDCLQILAY 296

RESULT 9
ID P79321 PRELIMINARY; PRT; 117 AA.
AC P79321;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE BETA CATENIN (FRAGMENT).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SMALL INTESTINE;
RA Winteroe A.K., Fredholm M.;
RT "Evaluation and characterization of a porcine small intestine cDNA
RT library.";
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z84131; CAB06327.1; -.
DR HSSP; P35222; IG3J.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 2.
DR PROSITE; PS50176; ARM_REPEAT; 2.
FT NON_TER 1
FT NON_TER 117
SQ SEQUENCE 117 AA; 12717 MW; 73F7CCF3B917A41A CRC64;

Query Match 77.3%; Score 157; DB 6; Length 117;
Best Local Similarity 100.0%; Pred. No. 2.3e-14;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 LAGGLQKMWVLLNKNVNFKLAITTDCLQILAY 41
Db 1 LAGGLQKMWVLLNKNVNFKLAITTDCLQILAY 32

RESULT 10
ID Q9PVF7 PRELIMINARY; PRT; 729 AA.
AC Q9PVF7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CELL-ADHESION PROTEIN PLAKOGLOBIN.
GN JUP.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-99386700; PubMed-10456847;
RA Cerda J., Reidenbach S., Pratzel S., Franke W.W.;
RT "Cadherin-catenin complexes during zebrafish oogenesis: heterotypic
RT junctions between oocytes and follicle cells.";
RL Biol. Reprod. 61:692-704(1999).
DR EMBL; AF099738; AAD36592.1; -.
DR HSSP; Q02248; IDOW.
DR ZFIN; ZDB-GENE-991207-22; jup.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 11.
DR SMART; SM00185; ARM; 8.
DR PROSITE; PS50176; ARM_REPEAT; 8.
SQ SEQUENCE 729 AA; 80033 MW; 91E00417B4FD8CEE CRC64;
```

```

Query Match 77.1%; Score 156.5; DB 13; Length 729;
Best Local Similarity 78.6%; Pred. No. 2.2e-13;
Matches 33; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 1 HOEGAM-AVRLAGGLQKMWVLLNKNVNFKLAITTDCLQILAY 41
Db 245 HOEGAKMAVRLADGLQRMVPLLKSNPKFLAITTDCLQLLSY 286

RESULT 11
ID O76152 PRELIMINARY; PRT; 773 AA.
AC O76152;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE BETA-CATENIN.
OS Ciona savignyi.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;
OC Clonidae; Clona.
OX NCBI_TaxID=51511;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-98443204; PubMed-9769178;
RA Yoshida S., Marikawa Y., Satoh N.;
RT "Regulation of the trunk-tail patterning in the ascidian embryo: a
RT possible interaction of cascades between lithium/beta-catenin and
RT localized maternal factor pem.";
RL Dev. Biol. 202:264-279(1998).
DR EMBL; AB012160; BAA32789.1; -.
DR HSSP; P35222; IG3J.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 11.
DR SMART; SM00185; ARM; 9.
DR PROSITE; PS50176; ARM_REPEAT; 7.
SQ SEQUENCE 773 AA; 85217 MW; C1340CF82AFEBDAB CRC64;
```

```

Query Match 77.1%; Score 156.5; DB 5; Length 773;
Best Local Similarity 83.3%; Pred. No. 2.3e-13;
Matches 35; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 1 HOEGAM-AVRLAGGLQKMWVLLNKNVNFKLAITTDCLQILAY 41
Db 254 HOEGAKEAVRLAGGLQKMWVLLSRDNRVNFKLAITTDCLQILAY 295

RESULT 12
ID Q25100 PRELIMINARY; PRT; 806 AA.
AC Q25100;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
```

```

Query Match      27.6%; Score 56; DB 5; Length 959;
Best Local Similarity 34.3%; Pred. No. 38;
Matches 12; Conservative 6; Mismatches 17; Indels 0; Gaps 0;

QY      2 OEGAMAVRLAGGLQKMWALLNKTNVKFLAITDCL 36
      : : : : : : : : : : : : : : : :
Ddb      39 KGVIFRTAGGVHIMVRLITKINEXIMEVVISIL 73

RESULT: 15
Q9XL01 PRELIMINARY; PRT; 607 AA.
IID Q9XL01
AC Q9XL01;

```

DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE NADH DEHYDROGENASE SUBUNIT 5.
OS Aythya americana (Redhead).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Aythya.
OX NCBI_TaxID=30385;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98424473; PubMed=9751919;
RA Johnson K.P., Sorenson M.D.;
RT "Comparing molecular evolution in two mitochondrial protein coding
genes (cytochrome b and ND2) in the dabbling ducks (Tribe: Anatini).";
RL Mol. Phylogenet. Evol. 10:82-94(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98393707; PubMed=9724766;
RA Mindell D.P., Sorenson M.D., Dimcheff D.E.;
RT "Multiple independent origins of mitochondrial gene order in birds.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:10693-10697(1998).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20113882; PubMed=10649633;
RA Sorenson M.D., Cooper A., Paxinos E.E., Quinn T.W., James H.F.,
RA Olson S.L., Fleischer R.C.;
RT "Relationships of the extinct moa-nalos, flightless Hawaiian
waterfowl, based on ancient DNA.";
RL Proc. R. Soc. Lond., B, Biol. Sci. 266:2187-2193(1999).
RN [4]
RP SEQUENCE FROM N.A.
RA Mindell D.P., Sorenson M.D., Dimcheff D.E., Hasegawa M., Ast J.C.,
RA Yuri T.;
RT "Interordinal relationships of birds and other reptiles based on whole
mitochondrial genomes.";
RL Syst. Biol. 48:138-152(1999).
RN [5]
RP SEQUENCE FROM N.A.
RA Sorenson M.D., Dimcheff D.E., Ast J.C., Yuri T., Mindell D.P.;
RT "Complete mitochondrial DNA sequences for five birds and a turtle."
Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -|- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
DR EMBL: AF090337; AAD32262.1; -;
DR InterPro: IPR001750; Oxidored_q1.
DR InterPro: IPR001516; Oxidored_q1_N.
DR Pfam: PF00361; oxidored_q1; 1.
DR Pfam: PF00662; oxidored_q1_N; 1.
KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
SQ SEQUENCE 607 AA; 66508 MW; 1D37736075BE8A09 CRC64;

Query Match 27.3%; Score 55.5; DB 8; Length 607;
Best Local Similarity 40.0%; Pred. No. 26;
Matches 14; Conservative 2; Mismatches 8; Indels 11; Gaps 1;

OY 4 GAMAVRLAGGLQKMKVALLNKTNVKEFLAITDCLQI 38
DB 353 GEQDIRKMGGLQKM-----LPVTTSLTI 376

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 16:07:53 ; Search time 158.47 Seconds
(without alignments)
30.139 Million cell updates/sec

Title: US-09-641-104A-8

Perfect score: 212

Sequence: 1 GNOESKLIILASGQPALVN.....TTTYEKLWTTTSRLVKLSV 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11107396 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A_Geneseq_032802:*

- 1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	212	100.0	43	20	AAV33219
2	212	100.0	781	21	AAV33219
3	212	100.0	781	21	AAV33219
4	212	100.0	781	21	AAV33219
5	212	100.0	781	21	AAV33219
6	212	100.0	781	21	AAV33219
7	212	100.0	781	21	AAV33219
8	209	98.6	43	20	AAV33219
9	186	87.7	840	22	ABB65819
10	186	87.7	840	22	ABB65819
11	186	87.7	840	22	ABB65821

12	53.8	700	12	AAR11354	Cellular Receptor
13	25.9	957	22	AAB80374	Secreted protein e
14	25.9	1003	21	AAB24058	Human PRO290 prote
15	25.9	1003	22	AAB65162	Human PRO290 (UNQ2
16	25.9	1004	21	AAV66639	Membrane-bound pro
17	25.9	1041	22	AAB80418	Gene #4 associated
18	25.7	78	21	AGC35383	zeas may protein f
19	25.7	128	21	AAG35382	zeas may protein f
20	25.7	495	21	AAB28393	Human angioleptin
21	25.7	495	21	AAV78904	Human angioleptin
22	25.7	495	21	AAV78907	Angioleptin fusio
23	25.7	497	17	AAW01410	Angioleptin fusio
24	25.7	497	17	AAW01410	Human TIE-2 ligand
25	25.7	497	17	AAW46504	Human TIE-2 ligand
26	25.7	497	19	AAW47531	Human TIE-2 ligand
27	25.7	498	17	AAW01409	Amino acid sequenc
28	25.7	498	17	AAW46503	Human TIE-2 ligand
29	25.7	498	19	AAW47528	Human TIE-2 ligand
30	25.7	498	19	AAW47530	Amino acid sequenc
31	25.7	498	21	AAW28391	Amino acid sequenc
32	25.7	498	21	AAV78902	Human angioleptin
33	25.7	498	21	AAV78905	Human angioleptin
34	25.0	407	22	ABG02327	Novel human diagno
35	24.5	173	21	AB43798	Human cancer assoc
36	24.5	411	22	AAU56274	Propionibacterium
37	24.5	1763	22	AAW23507	Human EST encoded
38	24.5	2473	22	AAW78911	Human protein SEQ
39	24.5	2668	22	AAW78995	Human protein SEQ
40	23.8	221	21	AGC33367	zeas may protein f
41	23.8	243	21	AGC33366	zeas may protein f
42	23.8	290	21	AGC33365	zeas may protein f
43	23.8	337	22	AAG39071	C glutamicum prote
44	23.8	337	22	AAB79594	Corynebacterium ql
45	23.6	347	15	AAW48726	G-protein coupled
46	23.6	347	17	AAW02698	G-protein coupled

ALIGNMENTS

RESULT 1

AA13219
ID AAY33219 standard; peptide; 43 AA.
XX AC
XX AAY33219;
XX AC
XX 18-NOV-1999 (first entry)
XX Human beta-catenin protein armadillo repeat arm5 fragment.
XX Beta-catenin; human; armadillo repeat; treatment; human disease; LEF-1;
XX modulator; interaction domain; transcription factor; TCF-4; oncogenic;
XX tumor suppressor gene; conductin; E-cadherin; screening; carcinoma; APC;
XX Wnt signalling pathway; tumor development; anti-oncogenic; melanoma;
XX organ regeneration; tissue regeneration; hair growth.
XX Homo sapiens.
XX OS
XX DE19909251-A1.
XX PN
XX PD
XX PF
XX PG
XX PH
XX PI
XX PJ
XX PK
XX PL
XX PM
XX PN
XX PO
XX PP
XX PQ
XX PR
XX PS
XX PT
XX PU
XX PV
XX PW
XX PX
XX PY
XX PZ
XX QA
XX QB
XX QC
XX QD
XX QE
XX QF
XX QG
XX QH
XX QI
XX QJ
XX QK
XX QL
XX QM
XX QN
XX QO
XX QP
XX QQ
XX QR
XX QS
XX QT
XX QU
XX QV
XX QW
XX QX
XX QY
XX QZ
XX RA
XX RB
XX RC
XX RD
XX RE
XX RF
XX RG
XX RH
XX RI
XX RJ
XX RK
XX RL
XX RM
XX RN
XX RO
XX RP
XX RQ
XX RR
XX RS
XX RT
XX RU
XX RV
XX RW
XX RX
XX RY
XX RZ
XX SA
XX SB
XX SC
XX SD
XX SE
XX SF
XX SG
XX SH
XX SI
XX SJ
XX SK
XX SL
XX SM
XX SN
XX SO
XX SP
XX SQ
XX SR
XX SS
XX ST
XX SU
XX SV
XX SW
XX SX
XX SY
XX SZ
XX TA
XX TB
XX TC
XX TD
XX TE
XX TF
XX TG
XX TH
XX TI
XX TJ
XX TK
XX TL
XX TM
XX TN
XX TO
XX TP
XX TQ
XX TR
XX TS
XX TT
XX TU
XX TV
XX TW
XX TX
XX TY
XX TZ
XX UA
XX UB
XX UC
XX UD
XX UE
XX UF
XX UG
XX UH
XX UI
XX UJ
XX UK
XX UL
XX UM
XX UN
XX UO
XX UP
XX UQ
XX UR
XX US
XX UT
XX UV
XX UW
XX UX
XX UY
XX UZ
XX VA
XX VB
XX VC
XX VD
XX VE
XX VF
XX VG
XX VH
XX VI
XX VJ
XX VK
XX VL
XX VM
XX VN
XX VO
XX VP
XX VQ
XX VR
XX VS
XX VT
XX VU
XX VV
XX VW
XX VX
XX VY
XX VZ
XX WA
XX WB
XX WC
XX WD
XX WE
XX WF
XX WG
XX WH
XX WI
XX WJ
XX WK
XX WL
XX WM
XX WN
XX WO
XX WP
XX WQ
XX WR
XX WS
XX WT
XX WU
XX WV
XX WW
XX WX
XX WY
XX WZ
XX XA
XX XB
XX XC
XX XD
XX XE
XX XF
XX XG
XX XH
XX XI
XX XJ
XX XK
XX XL
XX XM
XX XN
XX XO
XX XP
XX XQ
XX XR
XX XS
XX XT
XX XU
XX XV
XX XW
XX XX
XX XY
XX XZ
XX YA
XX YB
XX YC
XX YD
XX YE
XX YF
XX YG
XX YH
XX YI
XX YJ
XX YK
XX YL
XX YM
XX YN
XX YO
XX YP
XX YQ
XX YR
XX YS
XX YT
XX YU
XX YV
XX YW
XX YX
XX YY
XX YZ
XX ZA
XX ZB
XX ZC
XX ZD
XX ZE
XX ZF
XX ZG
XX ZH
XX ZI
XX ZJ
XX ZK
XX ZL
XX ZM
XX ZN
XX ZO
XX ZP
XX ZQ
XX ZR
XX ZS
XX ZT
XX ZU
XX ZV
XX ZW
XX ZX
XX ZY
XX ZZ

Agents for treating human diseases, particularly cancer, modulate interaction of beta-catenin with transcription factors or tumor

PT suppressor gene products -

XX Disclosure; Page 7; 16pp; German.

XX This invention describes a novel agent (A) for treating human disease

CC which is based on substances (I) that modulate (inhibit or promote) the

CC interaction of beta-catenin with transcription factors or products of

CC tumor suppressor genes. The invention also describes (a) peptides (II)

CC comprising part of the LEF-1/TCF-4 transcription factors, or their

CC variants and mutants; (b) peptides and related molecules (III) from the

CC armadillo domain (arm units 3-8) of beta-catenin, and mutants of the

CC complete beta-catenin molecule, that include at least one of the

CC specific interaction domains for LEF-1, TCF-4, APC, conductin or

CC E-cadherin; and (c) enzyme-linked immunosorbent assay (ELISA) for

CC screening substance libraries for compounds that modulate interaction of

CC beta-catenin with LEF-1/TCF, APC, conductin or E-cadherin. Beta-catenin

CC is a key compound in the Wnt signalling pathway and is involved in

CC development of tumors. Generally its interaction with LEF-1 or TCF-4 is

CC oncogenic but interaction with APC, conductin or E-cadherin is

CC anti-oncogenic. (A) which inhibit interaction are particularly used to

CC treat tumors, especially carcinoma of the colon and melanoma, but also,

CC where they promote interaction, to stimulate regeneration of organs and

CC tissues, specifically hair growth. AAY33217-Y33222 represent human

CC beta-catenin armadillo repeat fragments described in the method of the

CC invention.

XX Sequence 43 AA;

SQ

Query Match 100.0%; Score 212; DB 20; Length 43;

Best Local Similarity 100.0%; Pred. No. 4.2e-24;

Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GNQESKLIILASGGPQALVNMRTYTYEKLWTSRVKLVLSV 43

Db 1 gnqesklililassggpqqalvnmrtytyekllwttsrvklvlsv 43

RESULT 2

AAB07290

ID AAB07290 standard; Protein; 781 AA.

XX AAB07290;

XX 01-NOV-2000 (first entry)

XX Human beta catenin.

DE Beta catenin; cadherin; metastasis; cadherin-associated protein;

KW human; colorectal cancer; melanoma; antisense oligonucleotide;

KW gene therapy.

XX Homo sapiens.

XX US6066500-A.

XX 23-MAY-2000.

XX 25-JUN-1999; 99US-0344519.

XX 25-JUN-1999; 99US-0344519.

XX (ISIS-) ISIS PHARM INC.

XX Bennett CF, Cowse LM;

XX WPI; 2000-410651/35.

DR N-PSDB; AAA58320.

XX New antisense compounds targeting nucleic acids encoding human beta

PT catenin (HBC) useful for treating diseases associated with HBC

PT expression and as prophylaxis to prevent or delay infection,

PT inflammation or tumor formation

XX Example 13; Columns 45-52; 35pp; English.

XX Beta catenin is a member of the catenin family of cytosolic proteins and

CC a key member of the Wnt signalling pathway. Catenins interact with the

CC cytoplasmic domains of cadherin glycoproteins, and are important in

CC maintaining cell adhesiveness. The loss of cell adhesiveness is

CC implicated in metastasis. Beta catenin is also known as

CC cadherin-associated protein and is implicated in colorectal cancer and

CC melanoma. The present sequence is the human beta catenin protein. The

CC coding sequence of this protein was used in the present invention to

CC design antisense oligonucleotides (AAA58327-A58366). The

CC oligonucleotides are capable of hybridising to human beta catenin, in

CC order to inhibit expression of human beta catenin. The oligonucleotides

CC may be used in gene therapy for colorectal cancer or melanoma.

XX Sequence 781 AA;

SQ

Query Match 100.0%; Score 212; DB 21; Length 781;

Best Local Similarity 100.0%; Pred. No. 1.8e-22;

Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GNQESKLIILASGGPQALVNMRTYTYEKLWTSRVKLVLSV 43

Db 307 gnqesklililassggpqqalvnmrtytyekllwttsrvklvlsv 349

RESULT 3

AAY70740

ID AAY70740 standard; protein; 781 AA.

XX AAY70740;

XX 24-JUL-2000 (first entry)

XX Human beta-catenin.

DE Human beta-catenin; Wnt antagonist; contraceptive; contraceptive vaccine;

KW oocyte development; female primate contraception; oocyte viability;

KW monoclonal antibody; Wnt signalling.

XX Homo sapiens.

XX WO200021555-A1.

XX 20-APR-2000.

XX 13-OCT-1999; 99WO-US23640.

XX 15-OCT-1998; 98US-0104355.

XX (HARD.) HARVARD COLLEGE.

XX McMahon AP, Parr BA, Vaino S;

XX WPI; 2000-317845/27.

XX Contraceptive composition for inhibiting oocyte development in a female

PT primate comprises a Wnt polypeptide antagonist

XX Example 3; Page 26; 57pp; English.

XX The patent discloses a method of female primate contraception comprising

CC administering an antagonist of a Wnt polypeptide, inhibiting oocyte

CC development. Wnt polypeptides are useful for promotive maturation of an

CC immature oocyte. Wnt polypeptides are also useful for increasing the

CC number of mature oocytes and to enhance oocyte viability. Soluble

CC fragments of Wnt polypeptides have the ability to inhibit Wnt signalling,

CC e.g., by blocking binding of a naturally-occurring Wnt protein to its

CC receptor. They may be used to generate monoclonal antibodies which can

CC inhibit oocyte development. The present sequence is the human

CC beta-catenin protein. Dominant negative mutants of beta-catenin lack one

xx
The invention relates to novel isolated human secreted polypeptides (I)
and polynucleotides (II). (I) and (II) are useful for treating
CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
CC involved in increasing haematopoiesis, stem cell survival, bone growth
CC and remodelling. (I), (II) and modulators of (I) are useful for
CC prophylaxis or treatment of one or more cancers. (II) is also useful for
CC creating transgenic animals useful for studying the in vivo activities

XX Disclosure; Page 23-24; 33pp; English.

XX The present invention relates to a method for in vitro expansion of

XX mammalian stem or progenitor cells, which comprising increasing the

XX intracellular concentration of beta-catenin in a progenitor or stem cell

XX in an in vitro culture medium for a period sufficient for the progenitor

XX or stem cell to divide. The number of cells having the functional

XX phenotype of the stem or progenitor cells is expanded. The method is

XX useful for culturing stem cells and progenitor cells, which retain their

XX pluripotential phenotype after expansion, in vitro. The expanded cell

XX populations are useful as a source of stem cells, e.g. to reconstitute

XX function in a host that is deficient in a particular cell lineage or

XX lineages. The expanded cell populations are also useful in

XX transplantation to restore haematopoietic function to autologous or

XX allogeneic recipients. The present sequence is human beta-catenin

XX protein. Beta-catenin is a pivotal player in the signalling pathway

XX initiated by Wnt proteins, which are mediators of several developmental

XX processes.

SQ Sequence 781 AA;

Query Match 100.0%; Score 212; DB 22; Length 781;

Best Local Similarity 100.0%; Pred. No. 1.8e-22;

Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQESKLIILASGGPQALVNMRTYTYEKLWTSRVLKVLVS 43

Db 307 gqesklililassggpqqalvnmrtytyekllwttsrvlkvlsv 349

|||||

RESULT 6

AAE06039

ID AAE06039 standard; Protein: 781 AA.

AC AAE06039;

XX

XX 25-SEP-2001 (first entry)

DT

XX Mouse beta-catenin protein #2.

DE

XX Mouse; stem cell culturing; progenitor cell; pluripotential phenotype;

KW transplantation; haematopoietic function; allogeneic recipient;

KW signalling pathway; beta-catenin.

KW

XX Mus musculus.

OS

XX WO200152649-A1.

PN

XX 26-JUL-2001.

PD

XX 17-JAN-2001; 2001WO-US01459.

PF

XX 18-JAN-2000; 2000US-0176786.

PR

XX (STRD) UNIV LELAND STANFORD JUNIOR.

PA

XX Reya T, Nusse R, Weissman IL;

PI

XX WPI; 2001-465328/50.

DR

XX N-PSDB; AAD11165.

DR

XX In vitro expansion of mammalian stem or progenitor cells, useful for

XX producing cells that retain their pluripotential phenotype after

PT expansion for use in transplantation, involves increasing the levels of

PT beta-catenin in the cell -

PT

XX Disclosure; Page 28-30; 33pp; English.

PS

XX The present invention relates to a method for in vitro expansion of

XX mammalian stem or progenitor cells, which comprising increasing the

XX intracellular concentration of beta-catenin in a progenitor or stem cell

XX

CC in an in vitro culture medium for a period sufficient for the progenitor

CC or stem cell to divide. The number of cells having the functional

CC phenotype of the stem or progenitor cells is expanded. The method is

CC useful for culturing stem cells and progenitor cells, which retain their

CC pluripotential phenotype after expansion, in vitro. The expanded cell

CC populations are useful as a source of stem cells, e.g. to reconstitute

CC function in a host that is deficient in a particular cell lineage or

CC lineages. The expanded cell populations are also useful in

CC transplantation to restore haematopoietic function to autologous or

CC allogeneic recipients. The present sequence is mouse beta-catenin

CC protein. Beta-catenin is a pivotal player in the signalling pathway

CC initiated by Wnt proteins, which are mediators of several developmental

CC processes.

XX

SQ Sequence 781 AA;

Query Match 100.0%; Score 212; DB 22; Length 781;

Best Local Similarity 100.0%; Pred. No. 1.8e-22;

Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQESKLIILASGGPQALVNMRTYTYEKLWTSRVLKVLVS 43

Db 307 gqesklililassggpqqalvnmrtytyekllwttsrvlkvlsv 349

|||||

RESULT 7

AAU28306

ID AAU28306 standard; Protein: 800 AA.

XX

XX AAU28306;

AC

XX 18-DEC-2001 (first entry)

DT

XX Novel human secretory protein, Seq ID No 663.

DE

XX Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;

KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;

KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;

KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;

KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;

KW gut protection; lung; liver fibrosis; immune deficiency; infection;

KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;

KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;

KW fertility; analgesic; pain; antigen.

KW

XX Homo sapiens.

OS

XX WO200166689-A2.

PN

XX 13-SEP-2001.

PD

XX 05-MAR-2001; 2001WO-US04942.

PF

XX 07-MAR-2000; 2000US-0519705.

PR

XX 19-MAY-2000; 2000US-0574454.

PR

XX 17-JUN-2000; 2000US-0596193.

PR

XX 14-JUL-2000; 2000US-0616847.

PR

XX 19-SEP-2000; 2000US-0665363.

PR

XX 20-OCT-2000; 2000US-0693267.

PR

XX (HYSE-) HYSEQ INC.

PA

XX Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;

PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;

PI

XX WPI; 2001-589934/66.

DR

XX N-PSDB; AAS45206.

DR

XX Novel polypeptides and nucleic acids obtained from cDNA libraries

PT prepared from various human tissues, for diagnosis and treatment of

PT cancer, neurological, inflammatory, and autoimmune disorders -

PT

XX

PS XX Example 2; SEQ ID No 663; 107pp; English.

CC The invention relates to novel isolated human secreted polypeptides (I)

CC and polynucleotides (II). (I) and (II) are useful for treating

CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,

CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is

CC involved in increasing haematopoiesis, stem cell survival, bone growth

CC and remodeling. (I), (II) and modulators of (II) are useful for

CC prophylaxis or treatment of one or more cancers. (II) is also useful for

CC creating transgenic animals useful for studying the in vivo activities of

CC the polypeptide as well as for studying modulators of the polypeptides.

CC (I) induces the proliferation of neural cells and regeneration of nerve

CC and brain tissue and is useful for the treatment of central and

CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,

CC Parkinson's disease, Huntington's disease, and amyotrophic lateral

CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic

CC activity, regulation of haematopoiesis and is useful for treating myeloid

CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia

CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve

CC tissue growth, and in tissue repair, healing of burns, incisions,

CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative

CC disorders, or periodontal disease. Furthermore, (I) is also useful for

CC gut protection or regeneration and treatment of lung or liver fibrosis,

CC reperfusion injury in various tissues, various immune deficiencies and

CC disorders including severe combined immunodeficiency (SCID), bacterial or

CC fungal infections, autoimmune disorders e.g. multiple sclerosis,

CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic

CC reactions and conditions, such as asthma or other respiratory problems.

CC In addition, (I) affects biorhythms or circadian cycles of rhythms,

CC fertility, metabolism, catabolism, anabolism, storage or elimination of

CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides

CC analgesic effects or other pain reducing effects, immunoglobulin like

CC activity and can act as an antigen in a vaccine composition to raise an

CC immune response. AAU28020-AAU28395 represent novel human secreted protein

CC amino acid sequences of the invention.

XX SQ Sequence 800 AA;

Query Match 100.0%; Score 212; DB 22; Length 800;
 Best Local Similarity 100.0%; Pred. No. 1.9e-22;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GNOESKLIILASGGPQALVNIIMRTYTYEKLWTTSRVLKVLVS 43
 |||||
 Db 320 gngeskliliasgpgqalvnmrttyekllwttsrvlklvs 362

RESULT 8
 AAY33234
 ID AAY33234 standard; peptide; 43 AA.
 XX
 AC AAY33234;

XX 18-NOV-1999 (first entry)

DE Human beta-catenin protein mutant armadillo repeat arm 5.

XX Beta-catenin; human; armadillo repeat; treatment; human disease; LEF-1;
 KW modulator; interaction domain; transcription factor; TCF-4; oncogenic;
 KW tumor suppressor gene; conductin; E-cadherin; screening; carcinoma; APC;
 KW wnt signalling pathway; tumor development; anti-oncogenic; melanoma;
 KW organ regeneration; tissue regeneration; hair growth; mutant.

XX Homo sapiens.
 OS Synthetic.

XX DEL90909251-A1.

XX 26-AUG-1999.

XX 22-FEB-1999; 99DE-1009251.

XX

PR 21-FEB-1998; 98DE-1007390.
 XX
 PA (DELB) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.

XX Birchmeier W, Von.Kries J;

XX WPI; 1999-470389/40.

XX Agents for treating human diseases, particularly cancer, modulate
 PT interaction of beta-catenin with transcription factors or tumor
 PT suppressor gene products
 XX
 XX Example 4; Fig 5; 16pp; German.

XX This invention describes a novel agent (A) for treating human disease
 CC which is based on substances (I) that modulate (inhibit or promote) the
 CC interaction of beta-catenin with transcription factors or products of
 CC tumor suppressor genes. The invention also describes (a) peptides (II)
 CC comprising part of the LEF-1/TCF-4 transcription factors, or their
 CC variants and mutants; (b) peptides and related molecules (III) from the
 CC armadillo domain (arm units 3-8) of beta-catenin, and mutants of the
 CC complete beta-catenin molecule, that include at least one of the
 CC specific interaction domains for LEF-1, TCF-4, APC, conductin or
 CC E-cadherin; and (c) enzyme-linked immunosorbent assay (ELISA) for
 CC screening substance libraries for compounds that modulate interaction of
 CC beta-catenin with LEF-1/TCF, APC, conductin or E-cadherin. Beta-catenin
 CC is a key compound in the wnt signalling pathway and is involved in
 CC development of tumors. Generally its interaction with LEF-1 or TCF-4 is
 CC oncogenic but interaction with APC, conductin or E-cadherin is
 CC anti-oncogenic. (A) which inhibit interaction are particularly used to
 CC treat tumors, especially carcinoma of the colon and melanoma, but also,
 CC where they promote interaction, to stimulate regeneration of organs and
 CC tissues, specifically hair growth. AAY3230-Y3241 represent mutant
 CC human-beta-catenin armadillo repeat fragments described in the method of
 CC the invention.

XX SQ Sequence 43 AA;

Query Match 98.6%; Score 209; DB 20; Length 43;
 Best Local Similarity 97.7%; Pred. No. 1.2e-23;
 Matches 42; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GNOESKLIILASGGPQALVNIIMRTYTYEKLWTTSRVLKVLVS 43
 |||||
 Db 1 gngeskliliasgpgqalvnmrttyekllwttsrvlrvls 43

RESULT 9

ABB60196

ID ABB60196 standard; Protein; 840 AA.

XX

AC ABB60196;

XX

DT 26-MAR-2002 (first entry)

XX

DE Drosophila melanogaster polypeptide SEQ ID NO 7380.

XX

KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

XX

OS Drosophila melanogaster.

XX

FN WO200171042-A2.

XX

PD 27-SEP-2001.

XX

PE 23-MAR-2001; 2001WO-US09231.

XX

PR 23-MAR-2000; 2000US-191637P.

XX

PA (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;
 XX WPI; 2001-656860/75.
 DR N-PSDB; ABL04299.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX Disclosure; SEQ ID NO 7380; 21pp + Sequence Listing; English.
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 840 AA;

Query Match 87.7%; Score 186; DB 22; Length 840;
 Best Local Similarity 88.4%; Pred. No. 1.4e-18;
 Matches 38; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 GNOESKLIILASGGPOALVNIMRTYTYEKLWTTTSRVKLVSV 43
 DB 312 gngesklililiasggpnelvrmsydyekllwttsrvklvsv 354

RESULT 10
 ABB65819
 ID ABB65819 standard; Protein; 840 AA.
 XX
 XX ABB65819;
 AC
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 24249.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 XX 23-MAR-2001; 2001WO-US09231.
 XX
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 XX (PEKE) PE CORP NY.
 XX
 XX Venter JC, Adams M, Li PWD, Myers EW;
 XX WPI; 2001-656860/75.
 DR N-PSDB; ABL09922.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX Disclosure; SEQ ID NO 24249; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 840 AA;

Query Match 87.7%; Score 186; DB 22; Length 840;
 Best Local Similarity 88.4%; Pred. No. 1.4e-18;
 Matches 38; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 GNOESKLIILASGGPOALVNIMRTYTYEKLWTTTSRVKLVSV 43
 DB 312 gngesklililiasggpnelvrmsydyekllwttsrvklvsv 354

RESULT 11
 ABB65821
 ID ABB65821 standard; Protein; 840 AA.
 XX
 XX ABB65821;
 AC
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 24255.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 XX 23-MAR-2001; 2001WO-US09231.
 XX
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 XX (PEKE) PE CORP NY.
 XX
 XX Venter JC, Adams M, Li PWD, Myers EW;
 XX WPI; 2001-656860/75.
 DR N-PSDB; ABL09924.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX Disclosure; SEQ ID NO 24255; 21pp + Sequence Listing; English.
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX

AC AAB24058;
 XX
 DT 29-JAN-2001 (first entry)
 XX
 DE Human PRO290 protein sequence SEQ ID NO:7.
 XX
 DE Human; tumour; diagnosis; neoplastic disease; neoplastic cell growth;
 KW proliferation; tumorigenesis; identification; cancer; cytostatic;
 KW neotropic; neuroprotective; antiinflammatory; immunosuppressive;
 KW immunostimulant; antiangiogenic; leukaemia; lymphoid malignancy;
 KW neuronal disorder; glial disorder; astrocytal disorder; angiogenic;
 KW hypothalamic disorder; glandular disorder; macrophagal disorder;
 KW epithelial disorder; stromal disorder; blastocoeleic disorder;
 KW inflammatory disorder; immunologic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200053755-A2.
 XX
 PD 14-SEP-2000.
 XX
 XX 06-JAN-2000; 2000WO-US00376.
 XX
 PR 08-MAR-1999; 99WO-US05028.
 PR 02-JUN-1999; 99WO-US12252.
 PR 23-JUN-1999; 99US-0141037.
 PR 07-JUL-1999; 99US-0143048.
 PR 26-JUL-1999; 99US-0145698.
 PR 30-NOV-1999; 99WO-US28313.
 PR 20-DEC-1999; 99WO-US30911.
 PR 05-JAN-2000; 2000WO-US00219.
 XX
 PA (GETH) GENENTECH INC.
 XX
 XX Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hillan KJ, Roy MA;
 PI Watanabe CK, Wood WI;
 DR WPI: 2000-572270/53.
 DR N-PSDB; AAC58366.
 XX
 XX Thirty PRO polynucleotides encoding PRO polypeptides, useful in the
 PT treatment, diagnosis and prevention of cancer -
 PS
 PS Claim 61; Fig 4; 286pp; English.
 XX
 CC The present invention describes an isolated antibody that binds to
 CC one of the human PRO proteins designated PRO212, PRO290, PRO341, PRO535,
 CC PRO619, PRO177, PRO809, PRO830, PRO848, PRO943, PRO1005, PRO1009,
 CC PRO1025, PRO1030, PRO1097, PRO1107, PRO1111, PRO1153, PRO1182, PRO1184,
 CC PRO1187, PRO1281, PRO23, PRO39, PRO834, PRO1317, PRO1710, PRO2094,
 CC PRO2145 OR PRO2198. PRO antagonists can be used to inhibit tumour cell
 CC growth. The PRO polypeptides and nucleotides are useful in the
 CC treatment, diagnosis and prevention of cancer. The antibodies and other
 CC anti-tumour compounds may be used to treat various conditions, including
 CC those characterised by overexpression and/or activation of the amplified
 CC PRO genes. Exemplary conditions or disorders to be treated with such
 CC antibodies and other compounds include benign or malignant tumours
 CC (e.g., renal, liver, kidney, bladder, breast, gastric, ovarian,
 CC colorectal, prostate, pancreatic, lung, vulva, thyroid, hepatic
 CC carcinomas, sarcomas, glioblastomas, and various head and neck tumours),
 CC leukaemias and lymphoid malignancies, other disorders such as neuronal,
 CC glial, astrocytal, hypothalamic and other glandular, macrophagal,
 CC epithelial, stromal and blastocoeleic disorders, and inflammatory,
 CC angiogenic and immunologic disorders. AAC58242 to AAC58366 represent PCR
 CC primers and hybridisation probes used in the isolation of the human PRO
 CC sequences. AAC58367 to AAC58396 and AAB24057 to AAB24089 represent human
 CC PRO polynucleotide and protein sequences given in the exemplification of
 CC the present invention.
 XX
 XX Sequence 1003 AA;

Query Match 25.9%; Score 55; DB 21; Length 1003;

Best Local Similarity: 28.6%; Pred. No. 47;
 Matches 12; Conservative 12; Mismatches 14; Indels 4; Gaps 1;
 QY 111 CNGESKLIILASGPQALVNIMRTYTYEKLWTTSRVKLVLS 42
 DB 95 gldgklivvagpse----vrssqfarklwrssrrisqvs 991
 RESULT 15
 AAB65162
 ID AAB65162 standard; Protein; 1003 AA.
 XX
 AC AAB65162;
 XX
 XX 02-APR-2001 (first entry)
 XX
 DE Human; PRO290 (UNQ253), protein sequence SEQ ID NO:33.
 XX
 KW Human; secreted and transmembrane protein; PRO; cytostatic;
 KW cell death; cancer; chromosomal mapping; gene mapping; tissue typing;
 KW diagnostic assay.
 XX
 OS Homo sapiens.
 XX
 PN WO200073454-A1.
 XX
 PD 07-DEC-2000.
 XX
 PF 30-MAR-2000; 2000WO-US08439.
 XX
 PR 02-JUN-1999; 99WO-US12252.
 PR 23-JUN-1999; 99US-0141037.
 PR 07-JUL-1999; 99US-0143048.
 PR 20-JUL-1999; 99US-0144758.
 PR 26-JUL-1999; 99US-0145698.
 PR 28-JUL-1999; 99US-0146222.
 PR 17-AUG-1999; 99US-0149396.
 PR 15-SEP-1999; 99WO-US21090.
 PR 15-SEP-1999; 99WO-US21547.
 PR 08-OCT-1999; 99US-0158663.
 PR 30-NOV-1999; 99WO-US28313.
 PR 01-DEC-1999; 99WO-US28301.
 PR 16-DEC-1999; 99WO-US30095.
 PR 20-DEC-1999; 99WO-US30911.
 PR 05-JAN-2000; 2000WO-US00219.
 PR 06-JAN-2000; 2000WO-US00376.
 PR 11-FEB-2000; 2000WO-US03565.
 PR 18-FEB-2000; 2000WO-US04341.
 PR 22-FEB-2000; 2000WO-US04414.
 PR 24-FEB-2000; 2000WO-US04914.
 PR 24-FEB-2000; 2000WO-US05004.
 PR 02-MAR-2000; 2000WO-US05841.
 PR 15-MAR-2000; 2000WO-US06884.
 PR 20-MAR-2000; 2000WO-US07377.
 XX
 PA (GETH,) GENENTECH, INC.
 XX
 XX Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
 PI Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ;
 PI Grimaldi CJ, Gurney AL, Kijavini LJ, Napier MA, Pan J, Paoni NF;
 PI Roy MA, Stewart RA, Tumas D, Watanabe CK, Williams PM, Wood WI;
 PI Zhang Z;
 XX
 DR WPI: 2001-032160/04.
 DR N-PSDB; AAF44100.
 XX
 XX PRO polynucleotides used to produce polypeptides used to target
 PT bioactive molecules such as toxins, radiolabels or antibodies, to
 PT specific cells, to cause targeted cell death -
 PS Claim 12; Fig 23; 935pp; English.
 XX
 CC The present invention describes human secreted and transmembrane PRO

CC proteins. The PRO proteins have cytostatic activity. The PRO proteins
 CC can be used for targeted delivery of bioactive molecules, such as
 CC toxins, radiolabels or antibodies, that cause cell death. PRO nucleotide
 CC sequences, and their fragments, can be used as hybridisation probes, in
 CC chromosomal and gene mapping, and in the generation of anti-sense RNA
 CC and DNA. They may also be used to produce transgenic animals which are
 CC used to develop and screen therapeutically useful reagents. The PRO
 CC nucleotide and protein sequence can be used for tissue typing and in
 CC treating cancer. Anti-PRO antibodies can be used in diagnostic assays.
 CC AAF44270 to AAF44470 represent PCR primers and hybridisation probes used
 CC in the isolation of human PRO sequences. AAF44087 to AAF44269 and
 CC AAB65154 to AAB65300 represent human PRO polynucleotide and protein
 CC sequences given in the exemplification of the present invention.

XX
 SQ Sequence 1003 AA;

Query Match 25.9%; Score 55; DB 22; Length 1003;
 Best Local Similarity 28.6%; Pred. No. 47;
 Matches 12; Conservative 12; Mismatches 14; Indels 4; Gaps 1;

QY 1 GNQESKLIILASGGPQALVNIMRTYTYEKLWTSRVLKVL 42
 Db 954 gldgklivvvaqpse----vrsqfarklwsrrisqvs 991

Search completed: July 29, 2002, 16:07:54
 Job time: 644 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 16:05:08 ; Search time 57.6 seconds
(without alignments)
18.234 Million cell updates/sec

Title: US-09-641-104A-8

Perfect score: 212

Sequence: 1 GNOESKLIILASGGPQALVN.....TYTEKLLWTTSRVLKVLV 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	59	27.8	457	US-08-740-223A-14	Sequence 14, Appl
2	54.5	25.7	478	US-08-740-223A-7	Sequence 7, Appl
3	54.5	25.7	490	US-08-740-223A-12	Sequence 12, Appl
4	54.5	25.7	491	US-08-740-223A-13	Sequence 13, Appl
5	54.5	25.7	495	US-08-740-223A-26	Sequence 26, Appl
6	54.5	25.7	495	US-09-351-457-5	Sequence 5, Appl
7	54.5	25.7	495	US-09-561-500-5	Sequence 5, Appl
8	54.5	25.7	495	US-09-561-108-5	Sequence 5, Appl
9	54.5	25.7	497	US-08-373-579-4	Sequence 4, Appl
10	54.5	25.7	497	US-08-418-595-4	Sequence 4, Appl
11	54.5	25.7	497	US-08-665-926-4	Sequence 4, Appl
12	54.5	25.7	497	US-08-348-492-4	Sequence 4, Appl
13	54.5	25.7	497	US-09-162-437-4	Sequence 4, Appl
14	54.5	25.7	497	US-08-740-223A-4	Sequence 4, Appl
15	54.5	25.7	498	US-08-373-579-2	Sequence 2, Appl
16	54.5	25.7	498	US-08-418-595-2	Sequence 2, Appl
17	54.5	25.7	498	US-08-665-926-2	Sequence 2, Appl
18	54.5	25.7	498	US-08-348-492-2	Sequence 2, Appl
19	54.5	25.7	498	US-09-162-437-2	Sequence 2, Appl
20	54.5	25.7	498	US-08-740-223A-2	Sequence 2, Appl
21	54.5	25.7	498	US-08-740-223A-20	Sequence 20, Appl
22	54.5	25.7	498	US-09-351-457-2	Sequence 2, Appl
23	54.5	25.7	498	US-09-561-500-2	Sequence 2, Appl
24	54.5	25.7	498	US-09-561-108-2	Sequence 2, Appl
25	52	24.5	3898	US-08-750-717-2	Sequence 2, Appl
26	51	24.1	620	US-08-982-785A-10	Sequence 10, Appl
27	50	23.6	347	US-08-118-270-47	Sequence 47, Appl

Sequence 47, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 17, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 1, Appl
Sequence 6, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 5, Appl
Sequence 8, Appl
Sequence 11, Appl
Sequence 21, Appl
Sequence 44, Appl

ALIGNMENTS

RESULT 1
US-08-740-223A-14
; Sequence 14, Application US/08740223A
; Patent No. 6265564
; GENERAL INFORMATION:
; APPLICANT: Davis, et al.
; TITLE OF INVENTION: Expressed Ligand - Vascular
; TITLE OF INVENTION: Intercellular Signalling Molecule
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.
; STREET: 777 Old Saw Mill Road
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA: US/08740, 223A
; FILING DATE: 25-OCT-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 60/022/999
; FILING DATE: 02-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Cobert, Robert J
; REGISTRATION NUMBER: 36,108
; REFERENCE/DOCKET NUMBER: REG 333
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 914-345-7400
; TELEFAX: 914-345-7721
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 497 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: mTL1
; LOCATION: 1..497
; OTHER INFORMATION: mouse TIE-2 ligand 1

Query Match 27.8%; Score 59; DB 4; Length 497;


```

RESULT      5
US-08-740-223A-26
: sequence 26, Application US/08740223A
: Patent No. 6265564
: GENERAL INFORMATION:
: APPLICANT: Davis, et al.
: TITLE OF INVENTION: Expressed Ligand
: TITLE OF INVENTION: Intercellular
: NUMBER OF SEQUENCES: 28
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Regeneron Pharmaceuticals
: STREET: 777 Old Saw Mill Road
: CITY: Tarrytown
: STATE: NY
: COUNTRY: USA
: ZIP: 10591
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSEQ Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/740,223
: FILING DATE: 25-OCT-1996
: CLASSIFICATION: 536
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: USSN 60/02223
: FILING DATE: 02-AUG-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Cobert, Robert J
: REGISTRATION NUMBER: 36,108
: REFERENCE/DOCKET NUMBER: REG 333
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 914-345-7400
: TELEFAX: 914-345-7721
: INFORMATION FOR SEQ ID NO: 26:
: SEQUENCE CHARACTERISTICS:

```

```

US-09-561-500-5      Sequence 5, Application US/09561500
      Patent No. 6342219
      GENERAL INFORMATION:
      APPLICANT: Philip E. Thorpe
      APPLICANT: Rolf A. Brekken
      TITLE OF INVENTION: ANTIBODY COMPOUNDS
      FILE REFERENCE: 4001.002500
      CURRENT APPLICATION NUMBER: US/09/04-28
      CURRENT FILING DATE: 2000-04-28
      PRIOR APPLICATION NUMBER: 60/131,
      PRIOR FILING DATE: 1999-04-28
      NUMBER OF SEQ ID NOS: 44
      SOFTWARE: PatentIn Ver. 2.0
      SEQ_ID NO 5
      LENGTH: 495
      TYPE: PRT
      ORGANISM: Homo sapiens
US-09-561-500-5

```


TELECOMMUNICATION INFORMATION:
 TELEPHONE: (914) 345-7400
 TELEFAX: (914) 345-7721
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 497 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-418-595-4

Query Match 25.7%; Score 54.5; DB 2; Length 497;
 Best Local Similarity 42.5%; Pred. No. 4.7;
 Matches 17; Conservative 5; Mismatches 11; Indels 7; Gaps 2;

Qy 2 NOESKLIIILASGGPOALVIMRTYTYEK-LLWTTTSRLVKV 40
 Db 154 NOTSRLEI-----QLLENSLSTYKLEKQLLQQTNEILKI 187

RESULT 11
 US-08-665-926-4
 ; Sequence 4, Application US/08665926
 ; Patent No. 5851797
 ; GENERAL INFORMATION:
 ; APPLICANT: Valenzuela et al.
 ; TITLE OF INVENTION: TIE LIGAND-3, METHODS OF MAKING AND USES
 ; TITLE OF INVENTION: THEREOF
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Regeneron Pharmaceuticals, Inc.
 ; STREET: 777 Old Saw Mill River Road
 ; CITY: Tarrytown
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10591-6707
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/665,926
 ; FILING DATE: 19-JUN-1996
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Robert J. Cobert
 ; REGISTRATION NUMBER: 36,108
 ; REFERENCE/DOCKET NUMBER: REG 330-H
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (914) 345-7400
 ; TELEFAX: (914) 345-2113
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 497 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-665-926-4

Query Match 25.7%; Score 54.5; DB 2; Length 497;
 Best Local Similarity 42.5%; Pred. No. 4.7;
 Matches 17; Conservative 5; Mismatches 11; Indels 7; Gaps 2;

Qy 2 NOESKLIIILASGGPOALVIMRTYTYEK-LLWTTTSRLVKV 40
 Db 154 NOTSRLEI-----QLLENSLSTYKLEKQLLQQTNEILKI 187

RESULT 12
 US-08-348-492-4

; Sequence 4, Application US/08348492
 ; Patent No. 5879672
 ; GENERAL INFORMATION:
 ; APPLICANT: Davis, et al.
 ; TITLE OF INVENTION: TIE-2 LIGAND, METHOD OF MAKING AND USES
 ; TITLE OF INVENTION: THEREOF
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Regeneron Pharmaceuticals, Inc.
 ; STREET: 777 Old Saw Mill River Road
 ; CITY: Tarrytown
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10591
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/348,492
 ; FILING DATE: 02-DEC-1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/330,261
 ; FILING DATE: 27-OCT-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/319,932
 ; FILING DATE: 07-OCT-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kempler, Gail M.
 ; REGISTRATION NUMBER: 32,143
 ; REFERENCE/DOCKET NUMBER: REG 330B
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (914) 347-7000
 ; TELEFAX: (914) 347-2113
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 497 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-348-492-4

Query Match 25.7%; Score 54.5; DB 2; Length 497;
 Best Local Similarity 42.5%; Pred. No. 4.7;
 Matches 17; Conservative 5; Mismatches 11; Indels 7; Gaps 2;

Qy 2 NOESKLIIILASGGPOALVIMRTYTYEK-LLWTTTSRLVKV 40
 Db 154 NOTSRLEI-----QLLENSLSTYKLEKQLLQQTNEILKI 187

RESULT 13
 US-09-162-437-4
 ; Sequence 4, Application US/09162437
 ; Patent No. 6166185
 ; GENERAL INFORMATION:
 ; APPLICANT: Davis, et al.
 ; TITLE OF INVENTION: TIE-2 LIGAND, METHOD OF MAKING AND USES
 ; TITLE OF INVENTION: THEREOF
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Regeneron Pharmaceuticals, Inc.
 ; STREET: 777 Old Saw Mill River Road
 ; CITY: Tarrytown
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10591
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible

PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 60/022/999
FILING DATE: 02-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Cobert, Robert J
REGISTRATION NUMBER: 36,108
REFERENCE/DOCKET NUMBER: REG 333
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 497 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
FEATURE:
NAME/KEY: Human TIE-2 ligand 1
LOCATION: 1..2146
OTHER INFORMATION: from T98G clone
US-08-740-223A-4

Query Match 4; 25.7%; Score 54.5; DB 4; Length 497;
Best Local Similarity 42.5%; Pred. No. 4.7;
Matches 17; Conservative 5; Mismatches 11; Indels 7; Gaps 2;

' 2 NQESKLIILASGGPQALVNIMRTYTYEK - LLWTTSRVLKV 40
 QY ||| : ||| : ||| : ||| : ||| :
 Dp 154 NOTSRLEI-----OLLENSLSTYKLEKOLLQOOTNEILKI 187

RESULT .15

US-08-373-579-2
; Sequence '2, Application US/08373579
; Patent No. 5650490

GENERAL INFORMATION:
APPLICANT: Davis, et al.
TITLE OF INVENTION: TIE-2 LIGAND, METHOD OF MAKING AND USES
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Regeneron Pharmaceuticals, Inc.
STREET: 777 Old Saw Mill River Road
CITY: Tarrytown
STATE: New York
COUNTRY: USA

16COT:ATTN

```

? COMPUTER READABLE FORM:
?
? MEDIUM TYPE: Floppy disk
?
? COMPUTER: IBM PC compatible
?
? OPERATING SYSTEM: PC-DOS/MS-DOS
?
? SOFTWARE: PatentIn Release #1.0, Version #1.30
?
? CURRENT APPLICATION DATA:
?   APPLICATION NUMBER: US/08/373,579
?   FILING DATE: 17-JAN-1995
?   CLASSIFICATION: 435
?
? PRIOR APPLICATION DATA:
?   APPLICATION NUMBER: US 08/353,503
?   FILING DATE: 09-DEC-1994
?
? PRIOR APPLICATION DATA:
?   APPLICATION NUMBER: US 08/348,492
?   FILING DATE: 02-DEC-1994
?
? PRIOR APPLICATION DATA:
?   APPLICATION NUMBER: US 08/330,261
?   FILING DATE: 27-OCT-1994
?
? PRIOR APPLICATION DATA:
?   APPLICATION NUMBER: US 08/319,932
?   FILING DATE: 07-OCT-1994
?
? ATTORNEY/AGENT INFORMATION:
?   NAME: Cobert, Robert J.

```

NAME: Cobert, Robert J.

Query Match 25.7%; Score 54.5; DB 4; Length 497;
Best Local Similarity 42.5%; Pred. No. 4.7;
Matches 17: Conservative 5; Mismatches 11; Indels

QY 2 NQESKLIILASGGPQALVNIMRTYTEK-LLWTTSRVLKV 40
 II I I I I I I I I : I I I I I I I I : I I I :

D6 154 NOTSRLEI-----OLLENSTYKLEKKOLOOTNEILKI 187

RESULT 14

US-08-740-223A-4
; Sequence 4, Application US/08740223A

```

: Patent No. 6265564
:
: GENERAL INFORMATION:
:
: APPLICANT: Davis, et al.
:
: TITLE OF INVENTION: Expressed Ligand - Vascular
:
: TITLE OF INVENTION: Intercellular Signalling Molecule
:
: NUMBER OF SEQUENCES: 28
:
: CORRESPONDENCE ADDRESS:
:
: ADDRESSEE: Regeneron Pharmaceuticals, Inc.
:
: STREET: 777 Old Saw Mill Road
:
: CITY: Tarrytown
:
: STATE: NY
:
: COUNTRY: USA
:
: ZIP: 10591
:
: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Diskette
:
: COMPUTER: IBM Compatible
:
: OPERATING SYSTEM: DOS
:
: SOFTWARE: FastSeq Version 2.0
:
: CURRENT APPLICATION DATA:
:
: APPLICATION NUMBER: US/08/740,223A
:
: FILING DATE: 25-OCT-1996
:
: CLASSIFICATION: 536

```

CLASSIFICATION: 536

```

; REGISTRATION NUMBER: 36,108
; REFERENCE/DOCKET NUMBER: REG 330-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914) 345-7400
; TELEFAX: (914) 345-7721
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 498 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-373-579-2

```

```

Query Match      25.7%   Score 54.5; DB 1; Length 498;
Best Local Similarity 42.5%; Pred. No. 4.7;
Matches 17; Conservative 5; Mismatches 11; Indels 7; Gaps 2;

```

```

Oy  2 NOESKLIILASGGPOALVNMITYYK-LLWTTTSRLKV 40
    |||||  |||  |||  |||  |||  |||  |||  |||
Db 154 NOTSRLEI-----QLLENSLSYKLEKQLQLQQTNEILKI 187

```

Search completed: July 29, 2002, 16:05:09
Job time: 479 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 16:09:11 ; Search time 70.63 Seconds
(without alignments)
58.500 Million cell updates/sec

Title: US-09-641-104A-8
Perfect score: 212
Sequence: 1 GQESKLLILASGGPQALVN.....TYYEKLWTSRVLKVLVS 43

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR71.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	212	100.0	781	2 A38973	beta-catenin - hum
2	212	100.0	781	2 S35091	beta-catenin - mou
3	208	98.1	781	2 S35099	beta-catenin - Afr
4	195	92.0	820	2 S33794	hypothetical prote
5	194	91.5	744	2 A32905	plakoglobin, desmo
6	190	89.6	621	2 S35092	plakoglobin - mous
7	190	89.6	817	2 S33793	hypothetical prote
8	186	87.7	843	2 T12689	armadillo segment
9	185	87.3	738	2 S35093	plakoglobin - Afr
10	174	82.1	806	2 JC4835	beta-catenin - Hyd
11	91	42.9	811	2 T43175	armadillo protein
12	78.5	37.0	678	2 T23341	beta-catenin - Cae
13	57.5	27.1	364	2 B72672	hypothetical prote
14	57	26.9	811	2 T37631	probable exocyst c
15	56.5	26.7	307	2 C98207	oligopeptide trans
16	56.5	26.7	307	2 AF3079	hypothetical prote
17	56	26.4	293	2 A12139	ABC transport syst
18	55	25.9	842	2 G90576	heat shock atp-dep
19	54.5	25.7	170	2 T28696	hypothetical prote
20	54	25.5	214	2 C69879	hypothetical prote
21	54	25.5	451	2 S49829	VPS9 protein - yea
22	52	24.5	222	2 A75406	hydrolyase - yea
23	52	24.5	405	2 D90166	hypothetical prote
24	52	24.5	721	2 C84732	probable homeodoma
25	52	24.5	782	2 T43277	host cell factor 1
26	52	24.5	3898	2 S57437	genome polyprotein
27	52	24.5	3898	2 S58295	polyprotein - hog
28	51.5	24.3	99	2 G75354	hypothetical prote
29	51.5	24.3	307	2 F95949	probable oligopept

30 51.5 24.3 761 2 S44807 F44B9.1 protein -
31 51 24.1 170 2 S74037 hypothetical prote
32 51 24.1 241 1 PYNVW polyhedrin - Bomby
33 51 24.1 447 2 E84155 hypothetical prote
34 51 24.1 557 2 T44843 glucose-6-phosphat
35 51 24.1 578 2 S50446 VAC8 protein - yea
36 51 24.1 591 2 G86445 hypothetical prote
37 51 24.1 2048 2 C84609 hypothetical prote
38 51 24.1 2110 2 H96803 unknown protein T5
39 50.5 23.8 1082 2 S64903 regulatory protein T5
40 50 23.6 437 2 S67156 26S proteasome reg
41 50 23.6 445 2 T43420 probable protein k
42 50 23.6 584 2 B90281 hypothetical prote
43 50 23.6 585 2 T50325 probable ubiquitin
44 49.5 23.3 385 2 A81926 hypothetical prote
45 49.5 23.3 411 2 B97340 competence-damage

ALIGNMENTS

RESULT 1
A38973
beta-catenin - human
C:Species: Homo sapiens (man)
C:Date: 26-Jan-1996 #sequence_revision 08-Feb-1996 #text_change 08-Oct-1999
C:Accession: A38973; S55356; S31988
R:Huelsken, J.; Birchmeier, W.; Behrens, J.
J. Cell Biol. 127, 2061-2069, 1994
A:Title: E-cadherin and APC compete for the interaction with beta-catenin and the cyt
A:Reference number: A38973; MUID:95105247
A:Accession: A38973
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-781 <HUE>
A:Cross-references: GB:219054; NID:g38519; PIDN:CAA79497.1; PID:g38520
A:Experimental source: placenta
R:Nollet, F.; Bex, G.; Molemans, F.; van Roy, F.
submitted to the EMBL Data Library, June 1995
A:Description: H. sapiens beta-catenin mRNA.
A:Reference number: S55356
A:Accession: S55356
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-781 <NOL>
A:Cross-references: EMBL:X87838; NID:q1154853; PIDN:CAA61107.1; PID:g860988
C:Comment: The cytoplasmic protein beta-catenin normally binds to a region in the cyt
microfilament network.
C:Comment: Cellular levels of beta-catenin are regulated in part by the adenomatous p
d cancerous cell growth.
C:Genetics:
A:Gene: GDB:CTNNB1; CTNNB
A:Cross-references: GDB:141922; OMIM:116806
A:Map position: 3p22-3p21.3
C:Keywords: apoptosis; carcinogenesis; cell adhesion; cytosol
F:151-676/Region: 40-residue repeats

Query Match 100.0%; Score 212; DB 2; Length 781;
Best Local Similarity 100.0%; Pred. No. 6.1e-21;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GQESKLLILASGGPQALVNIRTYTEKLWTSRVLKVLVS 43
|||||
Db 307 GQESKLLILASGGPQALVNIRTYTEKLWTSRVLKVLVS 349
|||||

RESULT 2
S35091
beta-catenin - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999
C:Accession: S35091

R;Butz, S.; Stappert, J.; Weissig, H.; Kemler, R.
 Science 257, 1142-1144, 1992
 A:Title: Plakoglobin and beta-catenin: distinct but closely related.
 A:Reference number: S35091; MUID:92376536
 A:Accession: S35091
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-781 <BUT>
 A:Cross-references: EMBL:M90364; NID:gl92141; PIDN:AAA37280.1; PID:gl92142
 C:Keywords: cytoskeleton

Query Match 100.0%; Score 212; DB 2; Length 781;
 Best Local Similarity 100.0%; Pred. No. 6.1e-21;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNOESKLIILASGGPQALVNIMRTYTYEKLWTTTSRLVKLVSV 43
 |||||
 DB 307 GNOESKLIILASGGPQALVNIMRTYTYEKLWTTTSRLVKLVSV 349

RESULT 3
 S35099
 beta-catenin - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 10-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 21-Jul-2000
 C:Accession: S35099
 R;McCreary, P.D.; Turck, C.W.; Gumbiner, B.
 Science 254, 1359-1361, 1991
 A:Title: A homolog of the armadillo protein in Drosophila (plakoglobin) associated with
 A:Reference number: S35099; MUID:92073903
 A:Accession: S35099
 A:Molecule type: mRNA
 A:Residues: 1-781 <MCC>
 A:Cross-references: GB:M77013; NID:g214020; PIDN:AAA49670.1; PID:g214021
 C:Keywords: cytoskeleton

Query Match 98.1%; Score 208; DB 2; Length 781;
 Best Local Similarity 97.7%; Pred. No. 2.2e-20;
 Matches 42; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNOESKLIILASGGPQALVNIMRTYTYEKLWTTTSRLVKLVSV 43
 |||||
 DB 307 GNOESKLIILASGGPQALVNIMRTYTYEKLWTTTSRLVKLVSV 349

RESULT 4
 S33794
 hypothetical protein - sea urchin (Tripneustes gratilla)
 C:Species: Tripneustes gratilla
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Aug-1997
 C:Accession: S33794
 R;Rosenthal, E.
 Biochim. Biophys. Acta 1173, 337-341, 1993
 A:Title: Identification of homologues to beta-catenin/plakoglobin/armadillo in two inver
 A:Reference number: S33793; MUID:93305730
 A:Accession: S33794
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-820 <ROS>

Query Match 92.0%; Score 195; DB 2; Length 820;
 Best Local Similarity 93.0%; Pred. No. 1.4e-18;
 Matches 40; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GNOESKLIILASGGPQALVNIMRTYTYEKLWTTTSRLVKLVSV 43
 |||||
 DB 323 GNOESKLIILASGGPQALVNIMRTYTYEKLWTTTSRLVKLVSV 365

RESULT 5

A32905
 plakoglobin - desmosomal - human
 C:Species: Homo sapiens (man)
 C:Date: 22-Nov-1989 #sequence_revision 22-Nov-1989 #text_change 29-Aug-1997
 C:Accession: A32905
 R;Frankel, W.W.; Goldschmidt, M.D.; Zimbelmann, R.; Mueller, H.M.; Schiller, D.L.; Cow
 Proc. Natl. Acad. Sci. U.S.A. 86, 4027-4031, 1989
 A:Title: Molecular cloning and amino acid sequence of human plakoglobin, the common
 A:Reference number: A32905; MUID:89264555
 A:Accession: A32905
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-744 <FRA>
 A:Cross-references: GB:M23410
 C:Genetics:
 A:Gene: GDB:JUP
 A:Cross-references: GDB:126565; OMIM:173325
 A:Map position: 7pter-7qter
 C:Keywords: cytoskeleton

Query Match 91.5%; Score 194; DB 2; Length 744;
 Best Local Similarity 90.7%; Pred. No. 1.7e-18;
 Matches 39; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GNOESKLIILASGGPQALVNIMRTYTYEKLWTTTSRLVKLVSV 43
 |||||
 DB 297 GNOESKLIILASGGPQALVNIMRTYTYEKLWTTTSRLVKLVSV 339

RESULT 6
 S35092
 plakoglobin - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 17-Mar-1999
 C:Accession: S35092
 R;Butz, S.; Stappert, J.; Weissig, H.; Kemler, R.
 Science 257, 1142-1144, 1992
 A:Title: Plakoglobin and beta-catenin: distinct but closely related.
 A:Reference number: S35091; MUID:92376536
 A:Accession: S35092
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-621 <BUT>
 A:Cross-references: EMBL:M90365
 C:Keywords: cytoskeleton

Query Match 89.6%; Score 190; DB 2; Length 621;
 Best Local Similarity 88.4%; Pred. No. 5.1e-18;
 Matches 38; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GNOESKLIILASGGPQALVNIMRTYTYEKLWTTTSRLVKLVSV 43
 |||||
 DB 174 GNOESKLIILASGGPQALVNIMRTYTYEKLWTTTSRLVKLVSV 216

RESULT 7
 S33793
 hypothetical protein - spoonworm (Urechis caupo)
 C:Species: Urechis caupo
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Aug-1997
 C:Accession: S33793
 R;Rosenthal, E.
 Biochim. Biophys. Acta 1173, 337-341, 1993
 A:Title: Identification of homologues to beta-catenin/plakoglobin/armadillo in two in
 A:Reference number: S33793; MUID:93305730
 A:Accession: S33793
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-817 <ROS>

Query Match 89.6%; Score 190; DB 2; Length 817;
Best Local Similarity 90.7%; Pred. No. 6.9e-18;
Matches 39; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GNOESKLIILASGGPOALVNIMRTTYTEKLLWTTTSRVKLVSV 43
DB 330 GNOESKLIILASGGPGLVIRIMRSYTYEKLWTTTSRVKLVSV 372

RESULT 8

Tl2689
Armadillo segment polarity protein - fruit fly (*Drosophila melanogaster*)
A:Alternate names: protein 86E4.6
C:Species: *Drosophila melanogaster*
C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 21-Jul-2000
C:Accession: Tl2689; A31861
R:Ferraz, C.; Vidal, S.; Brun, C.; Bucheton, A.; Demaille, J.G.
submitted to the EMBL Data Library, January 1998
A:Description: Sequencing the distal X chromosome of *Drosophila melanogaster*.
A:Reference number: Z17572
A:Accession: Tl2689
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-843 <FER>
A:Cross-references: EMBL:AL021106; NID:e1371406; PID:e1249776; PIDN:CAA15946.1
A:Experimental source: clone cosmid 63B12
R:Rigleman, B.; Wieschaus, E.; Schedl, P.
Genes Dev. 3, 96-113, 1989
A:Title: Molecular analysis of the Armadillo locus: uniformly distributed transcripts and
A:Reference number: A31861; MUID:89211895
A:Accession: A31861
A:Molecule type: DNA
A:Residues: 1-843 <RIG>
A:Cross-references: EMBL:X54468; NID:g7610; PIDN:CAA38350.1; PID:g7611
C:Genetics:
A:Gene: arm
A:Cross-references: FlyBase:FBgn0000117
A:Introns: 11/1; 557/3; 624/3; 715/3; 748/1
A:Note: 86E4.6
C:Keywords: cytoskeleton

Query Match 87.7%; Score 186; DB 2; Length 843;
Best Local Similarity 88.4%; Pred. No. 2.5e-17;
Matches 38; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 GNOESKLIILASGGPOALVNIMRTTYTEKLLWTTTSRVKLVSV 43
DB 315 GNOESKLIILASGGPNELVIRIMRSYDYEKLWTTTSRVKLVSV 357

RESULT 9

S35093
Plakoglobin - African clawed frog
C:Species: *Xenopus laevis* (African clawed frog)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jul-2000
C:Accession: S35093; S24636
R:Fouquet, B.; Zimbelmann, R.; Franke, W.W.
Differentiation 51, 187-194, 1992
A:Title: Identification of plakoglobin in oocytes and early embryos of *Xenopus laevis*.
A:Reference number: S35093; MUID:93093332
A:Accession: S35093
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-738 <FOU>
A:Cross-references: EMBL:M95593; NID:g214656; PIDN:AAA49931.1; PID:g214657
R:Demarais, A.A.
submitted to the EMBL Data Library, June 1992
A:Reference number: S24636
A:Accession: S24636
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 133-184, 'V', 186-225, 'T', 227-292 <DEM>

A:Cross-references: EMBL:X67078; NID:g65252; PID:g65253
C:Keywords: cytoskeleton

Query Match 87.3%; Score 185; DB 2; Length 738;
Best Local Similarity 86.0%; Pred. No. 3e-17;
Matches 37; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 GNOESKLIILASGGPOALVNIMRTTYTEKLLWTTTSRVKLVSV 43
DB 294 GNOESKLIILGNGGPGGLVQIMRNRYTEKLLWTTTSRVKLVSV 336

RESULT 10

JC4835
beta-catenin - Hydra magnipapillata
C:Species: *Hydra magnipapillata*
C>Date: 15-Aug-1996 #sequence_revision 15-Oct-1996 #text_change 21-Jul-2000
C:Accession: JC4835
R:Hobmayer, E.; Hatta, M.; Fischer, R.; Fujisawa, T.; Holstein, T.W.; Sugiyama, T.
Gene 172, 155-159, 1996
A:Title: Identification of a Hydra homologue of the beta-catenin/plakoglobin/armadillo
A:Reference number: JC4835; MUID:96257271
A:Accession: JC4835
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-806 <HOB>
A:Cross-references: GB:U36781; NID:g1407600; PIDN:AAC47137.1; PID:g1407601
C:Comment: This protein plays roles in cadherin-mediated cell adhesion and in signal
C:Genetics:
A:Gene: betaCtn

Query Match 82.1%; Score 174; DB 2; Length 805;
Best Local Similarity 83.7%; Pred. No. 1.1e-15;
Matches 36; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 GNOESKLIILASGGPOALVNIMRTTYTEKLLWTTTSRVKLVSV 43
DB 366 GNOESKLIILSSGGPAELVIRIMRSYTYEKLTYTCRVKLVSV 408

RESULT 11

T43175
Armadillo protein homolog BAR-1 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 18-Feb-2000
C:Accession: T43175; T28813
R:Eisenmann, D.M.; Maloof, J.N.; Simske, J.S.; Kenyon, C.; Kim, S.K.
submitted to the EMBL Data Library, May 1998
A:Description: The beta-catenin homolog BAR-1 and LET-60 Ras coordinately regulate the
A:Reference number: Z22328
A:Accession: T43175
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-811 <EIS>
R:Minx, M.
submitted to the EMBL Data Library, January 1996
A:Description: The sequence of *C. elegans* cosmid C54D1.
A:Reference number: Z20527
A:Accession: T28813
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-811 <MIN>
A:Cross-references: EMBL:U46673; PIDN:AAC48154.1; GSPDB:GN00028; CESP:C54D1.6
C:Experimental source: strain Bristol N2; clone C54D1
C:Genetics:
A:Gene: bar-1; C54D1.6
A:Map position: X
A:Introns: 23/2; 65/3; 97/3; 132/3; 181/3; 223/2; 261/3; 332/3; 414/3; 452/3; 500/3;

Query Match 42.9%; Score 91; DB 2; Length 811;
Best Local Similarity 41.5%; Pred. No. 0.00029;
Matches 17; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 2 NOESKLIILASGGPQALVNIMRTTYEKLWTSRVKLVLS 42
DB 290 NTEQIKFVKMGPGQKLLMLLQHRVYENLLWRTTQLLKTF 330

RESULT 12
T23341
beta-catenin - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
C:Accession: T23341; T42221
R:Harris, B.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19729
A:Accession: T23341
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-678 <WIL>
A:Cross-references: EMBL:Z81564; PIDN:CAB04572.1; GSPDB:GN00019; CESP:K05C4.6
A:Experimental source: clone K05C4
R:Costa, M.; Raich, W.; Agbunag, C.; Hardin, J.; Priess, J.R.
submitted to the EMBL Data Library, July 1997
A:Description: A putative catenin-cadherin system mediates morphogenesis of the C. elegans
A:Reference number: Z22085
A:Accession: T42221
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-678 <COS>
A:Cross-references: EMBL:AF016853; PIDN:AAB94552.1
C:Genetics:
A:Gene: hmp-2; CESP:K05C4.6
A:Map position: 1
A:Introns: 14/1; 235/1; 290/1; 620/3

Query Match 37.08; Score 78.5; DB 2; Length 678;
Best Local Similarity 45.08; Pred. No. 0.012;
Matches 18; Conservative 10; Mismatches 11; Indels 1; Gaps 1;

QY 5 SKLIILASGGPQALVNIMRTTYEKLWTSRVKLVLS 43
DB 239 SKIFLSLLGQILVLSILREISDRKLIYTVVRCIRSLV 278

RESULT 13
B72672
hypothetical protein APE0800 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: B72672
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; KDNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum
A:Reference number: A72450; MUID:99310339
A:Accession: B72672
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-364 <KAW>
A:Cross-references: DDBJ:AP000060; NID:g5104188; PIDN:BAAT79778.1; PID:g1043564; PID:g510
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE0800

Query Match 27.18; Score 57.5; DB 2; Length 364;
Best Local Similarity 39.58; Pred. No. 4.8;
Matches 15; Conservative 7; Mismatches 9; Indels 7; Gaps 1;

QY 1 GNQESKLIILASG-----GPOALVNIMRTTYEKL 31
DB 173 GSREDAIIAAGPWTGFSGLQALSMLRVYRCEALI 210

RESULT 14
T37631
probable exocyst complex component - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T37631
R:Brown, D.; Churcher, C.M.; Barrell, B.G.; Rejandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z21733
A:Accession: T37631
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-811 <BRO>
A:Cross-references: EMBL:Z99091; PIDN:CAB11769.2; GSPDB:GN00066; SPDB:SPAC13F5.06c
A:Experimental source: strain 972h; cosmid c13F5
C:Genetics:
A:Gene: SPDB:SPAC13F5.06c
A:Map position: 1
A:Introns: 21/3; 48/1; 726/2; 760/2

Query Match 26.9%; Score 57; DB 2; Length 811;
Best Local Similarity 34.4%; Pred. No. 14;
Matches 11; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 12 SGGPQALVNIMRTTYEKLWTSRVKLVLSV 43
DB 154 SGKPQTLMDFFRTNNHDKMLLCAQRTQLLAL 185

RESULT 15
C98207
oligopeptide transport system permease protein oppB AGR_L1206 [imported] - Agrobacter
C:Species: Agrobacterium tumefaciens
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C:Accession: C98207
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorillo, B.; Goldm
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A:Reference number: A97359; PMID:11743194
A:Accession: C98207
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-307 <KUR>
A:Cross-references: GB:AE007870; PIDN:AAK89181.1; PID:g15158997; GSPDB:GN00170
C:Genetics:
A:Gene: AGR_L1206
A:Map position: linear chromosome

Query Match 26.7%; Score 56.5; DB 2; Length 307;
Best Local Similarity 37.1%; Pred. No. 5.5;
Matches 13; Conservative 5; Mismatches 8; Indels 9; Gaps 1;

QY 7 LIILASGG-----PQALVNIMRTTYEKLW 32
DB 26 LMRPAGGPNLERPLPPOTMENLMRTYHLDPLW 60

Search completed: July 29, 2002, 16:09:13
Job time: 723 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 16:09:52 ; Search time 34.24 Seconds
(without alignments)
48.626 Million cell updates/sec

Title: US-09-641-104A-8
Perfect score: 212
Sequence: 1 GNQESKLIIILASGGPQALVN.....TYYEKLLWTTSRVLKLSV 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	212	100.0	781	1 CTNB_HUMAN	P35222 homo sapien
2	212	100.0	781	1 CTNB_MOUSE	Q02248 mus musculus
3	212	100.0	781	1 CTNB_RAT	Q9wu82 rattus norv
4	208	98.1	781	1 CTNB_XENLA	P26233 xenopus lae
5	195	92.0	820	1 CTNB_TRIGR	P35223 tripneustes
6	194	91.5	743	1 PLAK_HUMAN	P14923 homo sapien
7	190	89.6	621	1 PLAK_MOUSE	Q02257 mus musculus
8	190	89.6	818	1 CTNB_URECA	P35224 urechis cau
9	186	87.7	813	1 ARM_MUSDO	Q02453 musca domes
10	186	87.7	843	1 ARM_DROME	P18824 drosophila
11	185	87.3	738	1 PLAK_XENLA	P30998 xenopus lae
12	54.5	25.7	110	1 AGP1_RAT	O33460 rattus norv
13	54.5	25.7	481	1 AGP1_BOVIN	O18920 bos taurus
14	54.5	25.7	498	1 AGP1_HUMAN	Q15389 homo sapien
15	54.5	25.7	498	1 AGP1_MOUSE	O08538 mus musculus
16	54	25.5	451	1 VPS9_YEAST	P54787 saccharomyc
17	52.5	24.8	416	1 CHRA_PSEAE	P14285 pseudomonas
18	51.5	24.3	761	1 YL31_CAEEL	P34422 caenorhabdi
19	51	24.1	557	1 GFI1_ACILW	Q9rmcl acinetobact
20	51	24.1	578	1 VAC8_YEAST	P39968 saccharomyc
21	50.5	23.8	1082	1 RGR1_YEAST	P19263 saccharomyc
22	50	23.6	437	1 SUG2_YEAST	P53549 saccharomyc
23	50	23.6	445	1 MEK1_SCHPO	Q10292 schizosacch
24	49	23.1	875	1 TRAC_ECOLI	P18004 escherichia
25	48.5	22.9	952	1 UVR4_MYCGE	P47660 mycoplasma
26	48	22.6	86	1 RL3E_ARCFU	O28389 archaeoglob
27	48	22.6	276	1 UL34_HSV6U	P52465 human herpe
28	48	22.6	298	1 BLAC_PROMI	P30897 proteus mir
29	48	22.6	326	1 Y800_PYRAB	Q9uze8 pyrococcus
30	48	22.6	384	1 VATIC_ASCSS	Q9ndr5 ascidia syd
31	48	22.6	398	1 PRSX_CAEEL	O17071 caenorhabdi
32	48	22.6	430	1 ER24_ASCIM	P78575 ascobolus i
33	48	22.6	940	1 SVI_BUCAI	P57249 buchnera ap

RESULT

CTNB_HUMAN

ID CTNB_HUMAN STANDARD; PRT; 781 AA.

AC P35222;

DT 01-FEB-1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE Beta-catenin.

GN CTNNB1 OR CTNNB.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Placenta;

RX MEDLINE=95105247; PubMed=7806582;

RA Huelken J., Birchmeier W., Behrens J.;

RT "E-cadherin and APC compete for the interaction with beta-catenin and

the cytoskeleton.";

RL J. Cell Biol. 127:2061-2069(1994).

RN [2]

RP REVIEW.

RX MEDLINE=20145417; PubMed=10679188;

RA Kikuchi A.;

RT "Regulation of beta-catenin signaling in the Wnt pathway.";

RL Biochem. Biophys. Res. Commun. 268:243-248(2000).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 133-664.

RX MEDLINE=20578906; PubMed=11136974;

RA Graham T.A., Weaver C., Mao F., Kimelman D., Xu W.;

RT "Crystal structure of a beta-catenin/Tcf complex.";

RL Cell 103:885-896(2000).

CC !-FUNCTION: INVOLVED IN THE REGULATION OF CELL ADHESION AND IN

SIGNAL TRANSDUCTION THROUGH THE WNT PATHWAY.

CC !-SUBUNIT: THE CYTOSOLIC FORM BOUND TO AXIN ALLOWS THE STIMULATION

OF ITS DEGRADATION. THE CYTOSOLIC FORM ALSO INTERACTS WITH APC AND

E-CADHERIN. THE NUCLEAR FORM BINDS TCF/LEF-1 AND MAY ALSO BIND

PONTIN52 AND DUPLIN. CAN HETERODIMERIZE WITH ALPHA-CATENIN AND

ALSO FORM A TRIMERIC COMPLEX WITH CADHERIN. MEDIATES THE ANCHORAGE

OF THE CADHERIN COMPLEX TO THE CORTICAL ACTIN CYTOSKELETON.

CC !-SUBCELLULAR LOCATION: CYTOPLASMIC WHEN IT IS UNSTABILIZED (HIGH

LEVEL OF PHOSPHORYLATION). TRANSLOCATES TO THE NUCLEUS WHEN IT IS

STABILIZED (LOW LEVEL OF PHOSPHORYLATION).

CC !-PTM: PROBABLY PHOSPHORYLATED BY GSK-3B. WNT-DEPENDENT ACTIVATION

OF DVL ANTAGONIZES THE ACTION OF GSK-3B. PHOSPHORYLATION OF BETA-

CATENIN IS REDUCED AND THE PROTEIN IS NO LONGER DEGRADED,

RESULTING IN ITS ACCUMULATION IN CYTOPLASM.

CC !-SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.

CC !-SIMILARITY: CONTAINS 12 ARM REPEATS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

Q15648 h peroxisom
Q07878 saccharomyc
P19712 hog cholel
O02739 bos taurus
P48595 homo sapien
P57448 buchnera ap
P38206 saccharomyc
P56518 strongyloc
O51578 borrelia bu
P40979 caldicellul
P38616 saccharomyc
Q46654 erwinia amy

ALIGNMENTS

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; X87838; CAA61107.1; -;
 CC EMBL; Z19054; CAA79497.1; -;
 CC PIR; S31988; S31988.
 CC PDB; 1G3J; 17-JAN-01.
 CC TRANSFAC; T02872; -;
 CC MIM; 116806; -;
 CC InterPro; IPR000225; Armadillo.
 CC Pfam; PF00514; Armadillo_seg; 12.
 CC SMART; SM00185; ARM; 11.
 CC PROSITE; PS50176; ARM_REPEAT; 9.
 CC Cell adhesion; Cytoskeleton; Structural protein; Nuclear protein;
 KW Repeat; 3D-structure.
 FT REPEAT 151 191 ARM 1.
 FT REPEAT 193 234 ARM 2.
 FT REPEAT 235 276 ARM 3.
 FT REPEAT 277 318 ARM 4.
 FT REPEAT 319 360 ARM 5.
 FT REPEAT 361 389 ARM 6.
 FT REPEAT 400 441 ARM 7.
 FT REPEAT 442 484 ARM 8.
 FT REPEAT 489 530 ARM 9.
 FT REPEAT 531 571 ARM 10.
 FT REPEAT 594 636 ARM 11.
 FT REPEAT 637 666 ARM 12.
 SQ SEQUENCE 781 AA; 85496 MW; CB78F165A3EEF86E CRC64;

Query Match 100.0%; Score 212; DB 1; Length 781;
 Best Local Similarity 100.0%; Pred. No. 7.5e-22;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GN0ESKLIILASGGPQALVNMRTYTYEKLWTSRVKLVSV 43
 DB 307 GN0ESKLIILASGGPQALVNMRTYTYEKLWTSRVKLVSV 349

RESULT 2
 CTNB_MOUSE STANDARD; PRT; 781 AA.
 AC Q02248;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Beta-catenin.
 GN CTNNB1 OR CATNB.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP MEDLINE-92376536; PubMed=1509266;
 RA Butz S., Stappert J., Weissig H., Kemler R.;
 RT "Plakoglobin and beta-catenin: distinct but closely related.";
 RL Science 257:1142-1144(1992).
 RN [2]
 RP REVIEW.
 RX MEDLINE-20145417; PubMed=10679188;
 RA Kikuchi A.;
 RT "Regulation of beta-catenin signaling in the Wnt pathway.";
 RL Biochem. Biophys. Res. Commun. 268:243-248(2000).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 150-665.
 RX MEDLINE-97442350; PubMed=9298899;
 RA Huber A.H., Nelson W.J., Weiss W.I.;
 RT "Three-dimensional structure of the armadillo repeat region of beta-catenin";
 RL Cell 90:871-882(1997).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 118-149 IN COMPLEX WITH

RP CTNNAL.
 RX MEDLINE-20337986; PubMed=10882138;
 RA Pokutta S., Weiss W.I.;
 RT "Structure of the dimerization and beta-catenin-binding region of
 alpha-catenin";
 RL Mol. Cell 5:533-543(2000).
 CC FUNCTION: INVOLVED IN THE REGULATION OF CELL ADHESION AND IN
 CC SIGNAL TRANSDUCTION THROUGH THE WNT PATHWAY.
 CC SUBUNIT: THE CYTOSOLIC FORM BOUND TO AXIN ALLOWS THE STIMULATION
 CC OF ITS DEGRADATION. THE CYTOSOLIC FORM ALSO INTERACTS WITH APC AND
 CC E-CADHERIN. THE NUCLEAR FORM BINDS TCF/LEF-1 AND MAY ALSO BIND
 CC PONTIN2 AND DUPLIN. CAN HETERODIMERIZE WITH ALPHA-CATENIN AND
 CC ALSO FORM A TRIMERIC COMPLEX WITH CADHERIN. MEDIATES THE ANCHORAGE
 CC OF THE CADHERIN COMPLEX TO THE CORTICAL ACTIN CYTOSKELETON.
 CC SUBCELLULAR LOCATION: CYTOPLASMIC WHEN IT IS UNSTABILIZED (HIGH
 CC LEVEL OF PHOSPHORYLATION). TRANSLOCATES TO THE NUCLEUS WHEN IT IS
 CC STABILIZED (LOW LEVEL OF PHOSPHORYLATION).
 CC PTM: PROBABLY PHOSPHORYLATED BY GSK-3B. WNT-DEPENDENT ACTIVATION
 CC OF DVL ANTAGONIZES THE ACTION OF GSK-3B. PHOSPHORYLATION OF BETA-
 CC CATENIN IS REDUCED AND THE PROTEIN IS NO LONGER DEGRADED,
 CC RESULTING IN ITS ACCUMULATION IN CYTOPLASM.
 CC SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
 CC SIMILARITY: CONTAINS 12 ARM REPEATS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;
 or send an email to license@isb-sib.ch).

CC EMBL; M90364; AAA37280.1; -;
 CC PIR; S35091; S35091.
 CC PDB; 2BCT; 15-OCT-97.
 CC PDB; 3BCT; 19-NOV-97.
 CC PDB; 1DOW; 12-JUL-00.
 CC TRANSFAC; T02984; -;
 CC MGD; MGI:88276; Catnb.
 CC InterPro; IPR000225; Armadillo.
 CC Pfam; PF00514; Armadillo_seg; 12.
 CC SMART; SM00185; ARM; 11.
 CC PROSITE; PS50176; ARM_REPEAT; 9.
 CC Cell adhesion; Cytoskeleton; Structural protein; Nuclear protein;
 KW Repeat; 3D-structure.
 FT REPEAT 151 191 ARM 1.
 FT REPEAT 193 234 ARM 2.
 FT REPEAT 235 276 ARM 3.
 FT REPEAT 277 318 ARM 4.
 FT REPEAT 319 360 ARM 5.
 FT REPEAT 361 389 ARM 6.
 FT REPEAT 400 441 ARM 7.
 FT REPEAT 442 484 ARM 8.
 FT REPEAT 489 530 ARM 9.
 FT REPEAT 531 571 ARM 10.
 FT REPEAT 594 636 ARM 11.
 FT REPEAT 637 666 ARM 12.
 SQ SEQUENCE 781 AA; 85470 MW; D708F170A3FED6E CRC64;

Query Match 100.0%; Score 212; DB 1; Length 781;
 Best Local Similarity 100.0%; Pred. No. 7.5e-22;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GN0ESKLIILASGGPQALVNMRTYTYEKLWTSRVKLVSV 43
 DB 307 GN0ESKLIILASGGPQALVNMRTYTYEKLWTSRVKLVSV 349

RESULT 3
 CTNB_RAT STANDARD; PRT; 781 AA.
 ID CTNB_RAT
 AC Q9W082;

Db 307 GNQESKLIILASGGPOALVNMRTSYEKLWTSRVLKVLVS 349

RESULT 5

```
CTNB_TRIGR STANDARD; PRT; 820 AA.
AC P35223;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Beta-catenin.
OS Trineustes gratilla (Hawaiian sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinodermata; Echinodermata; Echinodermata; Echinodermata; Echinodermata;
OC Trineustes.
OC NCBI_TaxID=7673;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93305730; PubMed=8318544;
RA Rosenthal E.T.;
RT Identification of homologues to beta-catenin/plakoglobin/armadillo
RT in two invertebrates, Urechis caupo and Trineustes gratilla.
RL Biochim. Biophys. Acta 1173:337-341(1993).
CC -1- FUNCTION: BINDS TO THE CYTOPLASMIC DOMAIN OF THE CELL-CELL
CC ADHESION MOLECULE E-CADHERIN, AND PERHAPS TO OTHER (MEMBRANE)
CC PROTEINS. THE ASSOCIATION OF CATENINS TO CADHERINS PRODUCES A
CC COMPLEX WHICH IS LINKED TO THE ACTIN FILAMENT NETWORK, AND WHICH
CC SEEMS TO BE OF PRIMARY IMPORTANCE FOR CADHERINS CELL-ADHESION
CC PROPERTIES.
CC -1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -1- SIMILARITY: CONTAINS 10 ARM REPEATS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC of the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; L10354; AAA30089.1;
CC PIR; S33794; S33794.
CC HSP; Q02248; IDOW.
CC InterPro: IPR000225; Armadillo.
CC Pfam: PF00514; Armadillo_seg; 12.
CC SMART; SM00185; ARM; 12.
CC SMART; SM00185; ARM; 12.
CC PROSITE; PS50176; ARM_REPEAT; 9.
CC Cell adhesion; Cytoskeleton; Structural protein; Repeat.
CC REPEAT 157 196
CC REPEAT 199 239 ARM 1.
CC REPEAT 241 280 ARM 2.
CC REPEAT 283 322 ARM 3.
CC REPEAT 367 405 ARM 4.
CC REPEAT 406 445 ARM 5.
CC REPEAT 448 489 ARM 6.
CC REPEAT 495 535 ARM 7.
CC REPEAT 603 642 ARM 8.
CC REPEAT 644 683 ARM 9.
CC REPEAT 743 784 ARM 10.
CC SEQUENCE 820 AA; 89361 MW; 57255E0F57795FD3 CRC64;
```

Query Match 92.0%; Score 195; DB 1; Length 820;
Best Local Similarity 93.0%; Pred. No. 1.9e-19;
Matches 40; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GNQESKLIILASGGPOALVNMRTSYEKLWTSRVLKVLVS 43
|||||
Db 323 GNQESKLIILASGGPOALVNMRTSYEKLWTSRVLKVLVS 365
|||||

RESULT 6

ID PLAK_HUMAN STANDARD; PRT; 743 AA.

AC P14923;

```
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Junction plakoglobin (Desmoplakin II).
DE JUN OR DF3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8926455; PubMed=2726765;
RA Franke W.W.; Goldschmidt M.D.; Zimbelmann R.; Mueller H.M.;
RA Schiller D.L.; Cowin P.;
RT "Molecular cloning and amino acid sequence of human plakoglobin, the
RT common junctional plaque protein."
RL Proc. Natl. Acad. Sci. U.S.A. 86:4027-4031(1989).
CC -1- FUNCTION: COMMON JUNCTIONAL PLAQUE PROTEIN. THE MEMBRANE-
CC ASSOCIATED PLAQUES ARE ARCHITECTURAL ELEMENTS IN AN IMPORTANT
CC STRATEGIC POSITION TO INFLUENCE THE ARRANGEMENT AND FUNCTION
CC OF BOTH THE CYTOSKELETON AND THE CELLS WITHIN THE TISSUE. THE
CC PRESENCE OF PLAQUEGLUBIN IN BOTH THE DESMOSOMES AND IN THE
CC INTERMEDIATE JUNCTIONS SUGGESTS THAT IT PLAYS A CENTRAL ROLE IN
CC THE STRUCTURE AND FUNCTION OF SUBMEMBRANOUS PLAQUES.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC IN A SOLUBLE AND MEMBRANE-
CC ASSOCIATED FORM.
CC -1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -1- SIMILARITY: CONTAINS 8 ARM REPEATS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC of the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; M23410; AAA64895.1;
CC PIR; A32905; A32905.
CC HSP; Q02248; 2BCT.
CC MIM; 173325;
CC InterPro: IPR000225; Armadillo.
CC Pfam: PF00514; Armadillo_seg; 10.
CC SMART; SM00185; ARM; 8.
CC PROSITE; PS50176; ARM_REPEAT; 8.
CC Cell adhesion; Cytoskeleton; Structural protein; Repeat.
CC INIT_MET 0 0
CC REPEAT 141 179 ARM 1.
CC REPEAT 183 224 ARM 2.
CC REPEAT 225 252 ARM 3.
CC REPEAT 308 351 ARM 4.
CC REPEAT 389 430 ARM 5.
CC REPEAT 431 473 ARM 6.
CC REPEAT 478 521 ARM 7.
CC REPEAT 582 624 ARM 8.
CC REPEAT 624 664 ARM 9.
CC REPEAT 664 703 ARM 10.
CC CONFLICT 96 99
CC CONFLICT 139 139
CC CONFLICT 139 139
CC SEQUENCE 743 AA; 81498 MW; 472741F400D388FD CRC64;
```

Query Match 91.5%; Score 194; DB 1; Length 743;
Best Local Similarity 90.7%; Pred. No. 2.4e-19;
Matches 39; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GNQESKLIILASGGPOALVNMRTSYEKLWTSRVLKVLVS 43
|||||
Db 296 GNQESKLIILASGGPOALVNMRTSYEKLWTSRVLKVLVS 338
|||||

RESULT 7;

PLAK_MOUSE

```

ID AC      PLAK_MOUSE      STANDARD;      PRT;      621 AA.
DT Q02257;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Junction plakoglobin (Desmoplakin III) (Fragment).
GN JUP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
EN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92376536; PubMed=1509266;
RA Butz S.; Stappert J.; Weissig H.; Kemler R.;
RT "Plakoglobin and beta-catenin: distinct but closely related.";
RL Science 257:1142-1144(1992).
RN [2]
RP REVISIONS TO 294 AND 296.
RA Butz S.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: COMMON JUNCTIONAL PLAQUE PROTEIN. THE MEMBRANE-
CC ASSOCIATED PLAQUES ARE ARCHITECTURAL ELEMENTS IN AN IMPORTANT
CC STRATEGIC POSITION TO INFLUENCE THE ARRANGEMENT AND FUNCTION
CC OF BOTH THE CYTOSKELETON AND THE CELLS WITHIN THE TISSUE. THE
CC PRESENCE OF PLAKOGLOBIN IN BOTH THE DESMOSOMES AND IN THE
CC INTERMEDIATE JUNCTIONS SUGGESTS THAT IT PLAYS A CENTRAL ROLE IN
CC THE STRUCTURE AND FUNCTION OF SUBMEMBRANOUS PLAQUES.
CC -!- SUBUNIT: HOMODIMER.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC IN A SOLUBLE AND MEMBRANE-
CC ASSOCIATED FORM.
CC -!- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -!- SIMILARITY: CONTAINS AT LEAST 9 ARM REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M90365; AAB02885.1; -
DR PIR; S35092; S35092.
DR HSSP; Q02248; 1DOW.
DR MGD; MGI:96650; Jup.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 11.
DR SMART; SM00185; ARM; 8.
DR PROSITE; PS50176; ARM_REPEAT; 9.
DR Cell adhesion; Cytoskeleton; Structural protein; Repeat.
FT NON_TER      1      47      ARM 1.
FT REPEAT      8      47      ARM 2.
FT REPEAT     92     131      ARM 3.
FT REPEAT     134     173      ARM 4.
FT REPEAT     218     257      ARM 5.
FT REPEAT     259     296      ARM 6.
FT REPEAT     299     340      ARM 7.
FT REPEAT     346     386      ARM 8.
FT REPEAT     388     427      ARM 9.
FT REPEAT     450     489
SQ SEQUENCE     621 AA; 68111 MW; 17CF444607422BAA CRC64;

Query Match      89.6%; Score 190; DB 1; Length 621;
Best Local Similarity 88.4%; Pred. No. 7e-19;
Matches 38; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY      1 GNOESKLIILASGGPQALVNIMRTYTYEKLWTTTSRLKVLVS 43
DB      174 GNOESKLIILANGPGGLVIMRNYSYKLLWTTTSRLKVLVS 216

```

```

RESULT      8
CTNB_URECA
ID CTNB_URECA      STANDARD;      PRT;      818 AA.
AC P35224;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Beta-catenin.
DE Beta-catenin.
OS Urechis caupo (Inkeeper worm) (Spoonworm).
OC Eukaryota; Metazoa; Echiura; Xenopneusta; Urechidae; Urechis.
OX NCBI_TaxID=6431;
EN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93305730; PubMed=8318544;
RA Rosenthal E.T.;
RT "Identification of homologues to beta-catenin/plakoglobin/armadillo
RT in two invertebrates, Urechis caupo and Tripneustes gratilla.";
RL Biochim. Biophys. Acta 1173:337-341(1993).
CC -!- FUNCTION: BINDS TO THE CYTOPLASMIC DOMAIN OF THE CELL-CELL
CC ADHESION MOLECULE E-CADHERIN, AND PERHAPS TO OTHER (MEMBRANE)
CC PROTEINS. THE ASSOCIATION OF CATENINS TO CADHERINS PRODUCES A
CC COMPLEX WHICH IS LINKED TO THE ACTIN FILAMENT NETWORK, AND WHICH
CC SEEMS TO BE OF PRIMARY IMPORTANCE FOR CADHERINS CELL-ADHESION
CC PROPERTIES.
CC -!- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -!- SIMILARITY: CONTAINS 7 ARM REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L10355; AAA30330.1; -
DR HSSP; Q02248; 2BCT.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 12.
DR SMART; SM00185; ARM; 11.
DR PROSITE; PS50176; ARM_REPEAT; 9.
DR Cell adhesion; Cytoskeleton; Structural protein; Repeat.
FT REPEAT     164     203      ARM 1.
FT REPEAT     248     287      ARM 2.
FT REPEAT     412     451      ARM 3.
FT REPEAT     454     495      ARM 4.
FT REPEAT     501     541      ARM 5.
FT REPEAT     543     582      ARM 6.
FT REPEAT     648     687      ARM 7.
SQ SEQUENCE     818 AA; 89070 MW; 1DF174BEEF745C1D CRC64;

Query Match      89.6%; Score 190; DB 1; Length 818;
Best Local Similarity 90.7%; Pred. No. 9.6e-19;
Matches 39; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY      1 GNOESKLIILASGGPQALVNIMRTYTYEKLWTTTSRLKVLVS 43
DB      330 GNOESKLIILASGGPGLVIMRNYSYKLLWTTTSRLKVLVS 372

RESULT      9
ARM_MUSDO
ID ARM_MUSDO      STANDARD;      PRT;      813 AA.
AC Q02453;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Armadillo segment polarity protein.
GN ARM.
OS Musca domestica (House fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

```

SEQUENCE FROM N.A. (CYTOPLASMIC ISOFORM).

RP STRAIN-OREGON-R; PubMed-2707602;
RC MEDLINE=89211895; PubMed-2707602;
RX Riggleman B., Wieschaus E., Schedl P.;
RA "Molecular analysis of the armadillo locus: uniformly distributed
RT transcripts and a protein with novel internal repeats are associated
RT with a Drosophila segment polarity gene.";
RL Genes Dev. 3:96-113(1989).

[2]
RN SEQUENCE FROM N.A. (FUNCTION, TISSUE SPECIFICITY, AND ALTERNATIVE
RP SPLICING.
RC TISSUE-Head;
RX MEDLINE=98298928; PubMed=9635189;
RA Loureiro J., Peifer M.;
RT "Roles of Armadillo, a Drosophila catenin, during central nervous
RT system development.";
RL Curr. Biol. 8:622-632(1998).

[3]
RN SEQUENCE FROM N.A. (CYTOPLASMIC ISOFORM).
RP STRAIN-BERKELEY;
RC MEDLINE=20196006; PubMed=10731132;
RX Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Anantades P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Branton R.C., Rogers Y.-H.C., Blaziej R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benton P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houch J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Talali M., Kalush F.A., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzhy D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassaman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs-R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).

[4]
RN SEQUENCE FROM N.A. (CYTOPLASMIC ISOFORM).
RP STRAIN-OREGON-R;
RC MEDLINE=20196011; PubMed=10731137;
RX Benos'P.V., Gatt M.K., Ashburner M., Murphy L., Demallies J., Cadieu E.,
RA Barrall B.G., Ferraz C., Vidal S., Brun C., Demallies J., Borokova D.,
RA Dreano S., Gloux S., Lelaure V., Mottier S., Gallibert F., Borokova D.,
RA Minana B., Kafatos F.C., Louis C., Siden-Kiamos I., Bolshakov S.,
RA Papadogiannis G., Spanos L., Cox S., Madueno E., de Pablos B.,
RA Mocadell J., Peter A., Schoettler P., Werner M., Mourikoti F.,
RA Beinert N., Dowe G., Schaefer U., Jaekle H., Bucheton A.,
RA Callister D.M., Campbell L.A., Darlamitsou A., Henderson N.S.,
RA McMillan P.J., Salles C., Tait E.A., Valenti P., Saunders R.D.C.,
RA Glover D.M.;

RT "From sequence to chromosome: the tip of the x chromosome of D.
 RL melanogaster.";
 RN Science 287:2220-2222(2000).
 RP [5]
 RX PHOSPHORYLATION.
 RA MEDLINE=95113174; PubMed=7529201;
 RT Pelfer M., Pai L.-M., Casey M.;
 RA "Phosphorylation of the Drosophila adherens junction protein
 Armadillo: roles for wingless signal and zeste-white 3 kinase.";
 RL Dev. Biol. 166:543-556(1994).
 CC -!- FUNCTION: NEURAL ISOFORM MAY ASSOCIATE WITH CADN AND PARTICIPATE
 CC IN THE TRANSMISSION OF DEVELOPMENTAL INFORMATION. CAN ASSOCIATE
 CC WITH ALPHA-CATENIN. CYTOPLASMIC ISOFORM ACCUMULATES THROUGH WG
 CC SIGNALING. ARM FUNCTION IN WG SIGNAL TRANSDUCTION IS REQUIRED
 CC EARLY IN DEVELOPMENT FOR DETERMINATION OF NEUROBLAST FATE. ARM AND
 CC ABL PROTEINS FUNCTION COOPERATIVELY AT ADHERENS JUNCTIONS IN BOTH
 CC THE CNS AND EPIDERMIS.
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC AND MEMBRANE-ASSOCIATED WITH THE
 CC INNER SURFACE OF CELL MEMBRANE.
 CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS: CYTOPLASMIC (SHOWN HERE) AND
 CC NEURAL; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- TISSUE SPECIFICITY: CYTOPLASMIC ISOFORM IS PREDOMINANT BEFORE GERM
 CC BAND RETRACTION, AFTER RETRACTION AND DURING LARVAL STAGES. IT IS
 CC FOUND IN HIGH LEVELS IN SPECIFIC CELLS ALONG THE CNS MIDLINE.
 CC NEURAL ISOFORM IS FIRST SEEN AFTER GERM BAND RETRACTION IN THE
 CC AXON TRACTS OF THE CNS, ALSO PRESENT IN AXONS DURING LARVAL STAGES
 CC AND ACCUMULATES IN THE MOTOR NEURONS OF THE SEGMENTAL AND
 CC INTERSEGMENTAL NERVES AS THEY EXIT THE CNS. BOTH ISOFORMS
 CC ACCUMULATE IN THE PNS.
 CC -!- DEVELOPMENTAL STAGE: PRESENT AT ALL STAGES, BUT REACHED THE
 CC HIGHEST LEVELS DURING EARLY TO MID-EMBRYOGENESIS.
 CC -!- PTM: PHOSPHORYLATED ON SER, THR AND TYR RESIDUES. LEVEL OF
 CC PHOSPHORYLATION VARIES BOTH DURING EMBRYONIC DEVELOPMENT AND FROM
 CC EMBRYONIC TISSUE TO TISSUE. SGG IS REQUIRED FOR PHOSPHORYLATION
 CC AND WG SIGNAL NEGATIVELY REGULATES ARM PHOSPHORYLATION.
 CC HYPOPHOSPHORYLATED FORM OF ARM INCREASES IN STEADY-STATE LEVELS.
 CC -!- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
 CC -!- SIMILARITY: CONTAINS 12.5 ARM REPEATS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X54468; CAA39350.1; -;
 CC EMBL; AF001213; AAB58731.1; -;
 CC EMBL; AE003422; AAF45688.1; ALT_INIT.
 CC EMBL; AL021106; CAA15946.1;
 CC EMBL; AL021086; CAA15946.1;
 CC EMBL; AL021086; CAA15946.1; JOINED.
 CC EMBL; AL021086; CAA15935.1;
 CC EMBL; AL021106; CAA15935.1; JOINED.
 CC PIR; A31861; A31861.
 CC HSP; Q02248; IDOW.
 CC TRANSFAC; T02977; -;
 CC FlyBase; FBgn0000117; arm.
 CC InterPro; IPR000225; Armadillo.
 CC Pfam; PF00514; Armadillo_seg; 12.
 CC SMART; SM00185; ARM; 11.
 CC PROSITE; PS00176; ARM_REPEAT; 9.
 CC Developmental protein; Segmentation polarity protein; Repeat;
 CC Cell adhesion; Cytoskeleton; Structural protein; Phosphorylation;
 CC Alternative splicing.
 CC
 CC FT DOMAIN 1 158 ASP/GLU-RICH (ACIDIC).
 CC REPEAT 159 200 ARM 1.
 CC REPEAT 201 242 ARM 2.
 CC REPEAT 243 284 ARM 3.
 CC REPEAT 285 326 ARM 4.
 CC REPEAT 327 368 ARM 5.
 CC REPEAT 369 410 ARM 6.
 CC REPEAT 411 449 ARM 7.

FT REPEAT 450 496 ARM 8.
 FT REPEAT 497 538 ARM 9.
 FT REPEAT 539 584 ARM 10.
 FT REPEAT 585 608 ARM 11.
 FT REPEAT 609 647 ARM 12.
 FT REPEAT 648 689 ARM 13 (INCOMPLETE).
 FT REPEAT 689 718 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 718 843 LGPEAYEGLYGQGPSPVSHSHGGAFFHQOQYDTLPIIDSMQ
 FT VARSPLIC 718 843 GLEISPPGGGAGGAGNGGAVGGAGSGGGGNGAIPPSGA
 FT PTPSPYMDMDYGEIDAGALNFDLDAMPTPPNDNNNLAAYD
 FT TDC -> ILYO (IN NEURAL ISOFORM).
 SQ SEQUENCE 843 AA; 91152 MW; 40DAD6FB83163049 CRC64;
 Query Match 87.7%; Score 186; DB 1; Length 843;
 Best Local Similarity 88.4%; Pred. No. 3.6e-18;
 Matches 38; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1. GNOESKLIILASGGPQALVNMRTTYTYEKLMTTSRLVKLSV 43
 DB 315. GNOESKLIILASGGPQALVNMRTTYTYEKLMTTSRLVKLSV 357
 RESULT 11;
 PLAK_XENLA
 ID PLAK_XENLA STANDARD; PRT: 738 AA.
 AC P30998;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Junction plakoglobin (Desmoplakin III).
 OS Xenopus laevis (African clawed frog).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 CC Xenopodinae; Xenopus.
 CC NCBI_TaxID=89355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93093332; PubMed=1459359;
 RA Fouquet B., Zimbelmann R., Franke W.W.;
 RT Identification of plakoglobin in oocytes and early embryos of
 RT Xenopus laevis; maternal expression of a gene encoding a junctional
 RT plaque protein.";
 RL Differentiation 51:187-194(1992).
 RN [2]
 RP SEQUENCE OF 133-292 FROM N.A.
 RX MEDLINE=93012479; PubMed=1397690;
 RA de Marais A.A., Moon R.T.;
 RT The armadillo homologs beta-catenin and plakoglobin are
 RT differentially expressed during early development of Xenopus
 RT laevis.";
 RL Dev. Biol. 153:337-346(1992).
 CC -!- FUNCTION: COMMON JUNCTIONAL PLAQUE PROTEIN. THE MEMBRANE-
 CC ASSOCIATED PLAQUES ARE ARCHITECTURAL ELEMENTS IN AN IMPORTANT
 CC STRATEGIC POSITION TO INFLUENCE THE ARRANGEMENT AND FUNCTION
 CC OF BOTH THE CYTOSKELETON AND THE CELLS WITHIN THE TISSUE. THE
 CC PRESENCE OF PLAKOGLOBIN IN BOTH THE DESMOSOMES AND IN THE
 CC INTERMEDIATE JUNCTIONS SUGGESTS THAT IT PLAYS A CENTRAL ROLE IN
 CC THE STRUCTURE AND FUNCTION OF SUBMEMBRANOUS PLAQUES.
 CC -!- SUBUNIT: HOMODIMER.
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC IN A SOLUBLE AND MEMBRANE-
 CC ASSOCIATED FORM.
 CC -!- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
 CC -!- SIMILARITY: CONTAINS 10 ARM REPEATS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

```

DR EMBL: M95593; AAA49931.1;
DR EMBL: X67078; CAA47463.1;
DR PIR: S24636; S24636.
DR PIR: S35093; S35093.
DR HSSP: Q02248; 2BCT.
DR InterPro: IPR000225; Armadillo.
DR Pfam: PF00514; Armadillo_seg; 11.
DR SMART: SM00185; ARM; 10.
DR PROSITE: PS00176; ARM_REPEAT; 9.
KW Cell adhesion; Cytoskeleton; Structural protein; Repeat.
FT REPEAT 128 167 ARM 1.
FT REPEAT 212 251 ARM 2.
FT REPEAT 254 293 ARM 3.
FT REPEAT 338 377 ARM 4.
FT REPEAT 379 416 ARM 5.
FT REPEAT 419 460 ARM 6.
FT REPEAT 466 506 ARM 7.
FT REPEAT 508 547 ARM 8.
FT REPEAT 570 609 ARM 9.
FT REPEAT 611 650 ARM 10.
FT CONFLICT 185 185 I -> V (IN REF. 2).
FT CONFLICT 226 226 A -> T (IN REF. 2).
FT SEQUENCE 738 AA; 81711 MW; 569DBE69D08BBC58 CRC64;

Query Match 87.3%; Score 185; DB 1; Length 738;
Best Local Similarity 86.0%; Pred. NO. 4.3e-18;
Matches 37; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 GNQESKLIILASGGPQALVNMITYYEKLLWTTSRVLKVLV 43
DB 294 GNQESKLIILNGGPGQLVQIMRNYEKLWTTSRVLKVLV 336

RESULT 12
AGPL_BOVIN
ID AGPL_BOVIN STANDARD; PRT; 110 AA.
AC 035460;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Angiopoietin-1 (ANG-1) (Fragment).
GN Angiopoietin-1
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Placenta;
RX MEDLINE=98451564; PubMed=9776732;
RA Mandriota S.J., Pepper M.S.;
RT "Regulation of angiopoietin-2 mRNA levels in bovine microvascular endothelial cells by cytokines and hypoxia.";
RL Circ. Res. 83:852-859(1998).
CC -1- FUNCTION: BINDS AND ACTIVATES TIE2 RECEPTOR BY INDUCING ITS TYROSINE PHOSPHORYLATION. IMPLICATED IN ENDOTHELIAL DEVELOPMENTAL PROCESSES LATER AND DISTINCT FROM THAT OF VEGF. APPEARS TO PLAY A CRUCIAL ROLE IN MEDIATING RECIPROCAL INTERACTIONS BETWEEN THE ENDOTHELIUM AND SURROUNDING MATRIX AND MESENCHYME. MEDIATES BLOOD VESSEL MATURATION/STABILITY. IT MAY PLAY AN IMPORTANT ROLE IN THE HEART EARLY DEVELOPMENT (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DEVELOPMENTAL STAGE: FOUND TO BE EXPRESSED THROUGHOUT THE OVARIAN CYCLE.
CC -1- SIMILARITY: CONTAINS 1 FIBRINOGEN C-TERMINAL DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC EMBL: AF030376; AAC78246.1;
CC HSSP: P02671; 1FZD.

```

```

DR InterPro: IPR002181; Fibrinogen_C.
KW Glycoprotein; Coiled coil.
FT NONTER 1 1
FT DOMAIN 63 109 COILED COIL (POTENTIAL).
FT CARBOHYD 2 2 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 32 32 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NONTER 110 110
FT SEQUENCE 110 AA; 12860 MW; 36D345DECB1E1845 CRC64;

Query Match 25.7%; Score 54.5; DB 1; Length 110;
Best Local Similarity 42.5%; Pred. NO. 0.94;
Matches 17; Conservative 5; Mismatches 11; Indels 7; Gaps 2;

QY 2 GNQESKLIILASGGPQALVNMITYYEKLLWTTSRVLKVLV 40
DB 64 NQTSRLKLEI-----QLLENSLSYKLEKQLLOQTNEILK 97

RESULT 13
AGPL_BOVIN
ID AGPL_BOVIN STANDARD; PRT; 481 AA.
AC 018920;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Angiopoietin-1 precursor (ANG-1) (Fragment).
GN Angiopoietin-1
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=99054348; PubMed=9840613;
RA Goede V., Schmidt T., Kimmina S., Kozian D., Augustin H.G.;
RT "Analysis of blood vessel maturation processes during cyclic ovarian angiogenesis.";
RL Lab. Invest. 78:1385-1394(1998).
RN [2]
RP SEQUENCE OF 91-200 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=98451564; PubMed=9776732;
RA Mandriota S.J., Pepper M.S.;
RT "Regulation of angiopoietin-2 mRNA levels in bovine microvascular endothelial cells by cytokines and hypoxia.";
RL Circ. Res. 83:852-859(1998).
CC -1- FUNCTION: BINDS AND ACTIVATES TIE2 RECEPTOR BY INDUCING ITS TYROSINE PHOSPHORYLATION. IMPLICATED IN ENDOTHELIAL DEVELOPMENTAL PROCESSES LATER AND DISTINCT FROM THAT OF VEGF. APPEARS TO PLAY A CRUCIAL ROLE IN MEDIATING RECIPROCAL INTERACTIONS BETWEEN THE ENDOTHELIUM AND SURROUNDING MATRIX AND MESENCHYME. MEDIATES BLOOD VESSEL MATURATION/STABILITY. IT MAY PLAY AN IMPORTANT ROLE IN THE HEART EARLY DEVELOPMENT (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DEVELOPMENTAL STAGE: FOUND TO BE EXPRESSED THROUGHOUT THE OVARIAN CYCLE.
CC -1- SIMILARITY: CONTAINS 1 FIBRINOGEN C-TERMINAL DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC EMBL: AF030373; AAC61872.1;
CC EMBL: AF032923; AAC78245.1;
CC HSSP: P02671; 1FZD.

```

```

DR InterPro: IPR002181; Fibrinogen_C.
DR Pfam: PF00147; fibrinogen_C; 1.
DR PROSITE; PS00514; FIBRINAG_C_DOMAIN; 1.
KW Glycoprotein; Coiled coil; Signal.
FT SIGNAL 1 15
FT CHAIN 16 >481
FT DOMAIN 153 261
FT DOMAIN 283 >481
FT CARBOHYD 92 92
FT CARBOHYD 122 122
FT CARBOHYD 154 154
FT CARBOHYD 243 243
FT CARBOHYD 294 294
FT NON_TER 481
SQ SEQUENCE 481 AA; 55556 MW; 8BEC9ED84FC2BB50 CRC64;

Query Match 25.7%; Score 54.5; DB 1; Length 481;
Best Local Similarity 42.5%; Pred. No. 5.3;
Matches 17; Conservative 5; Mismatches 11; Indels 7; Gaps 2;

QY 2 NOESKLILASGGQALVNMRTYTYEK-LLWTSRVLKV 40
Db 154 NQTSRLEI-----QLLENSLSTYKLEKQLLOQTNEILKI 187

RESULT 14
AGPL_HUMAN
ID AGPL_HUMAN STANDARD; PRT; 498 AA.
AC Q1538;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Angiopietin-1 precursor (ANG-1).
GN ANGPT1 OR KIAA0003.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A., AND VARIANT GLY-269 DEL.
TISSUE=Fetal lung;
RA Davis S., Aldrich T.H., Jones P.F., Acheson A., Compton D.L., Jain V.,
RA Ryan T.E., Bruno J., Radziejewski C., Maisonnier P.C.,
RA Yancopoulos G.D.;
FT "Isolation of angiopietin-1, a ligand for the TIE2 receptor, by
secretion-trap expression cloning.";
RL Cell 87:1161-1169(1996).
[2]
SEQUENCE FROM N.A.
Obara O., Nagase T., Kikuno R., Nomura N.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE OF 307-498 FROM N.A.
TISSUE=Bone marrow;
RX MEDLINE=96051387; PubMed=7584026;
RA Nomura N., Miyajima N., Suzuki T., Tanaka A., Kawarabayashi Y.,
RA Sato S., Nagase T., Seki N., Ishikawa K.-I., Tabata S.;
RT "Prediction of the coding sequences of unidentified human genes. I.
The coding sequences of 40 new genes (K1AA0001-K1AA0040) deduced by
analysis of randomly sampled cDNA clones from human immature myeloid
cell line KG-1.";
RL DNA Res. 1:27-35(1994).
CC -1- FUNCTION: BINDS AND ACTIVATES TIE2 RECEPTOR BY INDUCING ITS
TYROSINE PHOSPHORYLATION. IMPLICATED IN ENDOTHELIAL DEVELOPMENTAL
PROCESSES LATER AND DISTINCT FROM THAT OF VEGF. APPEARS TO PLAY A
CRUCIAL ROLE IN MEDIATING RECIPROCAL INTERACTIONS BETWEEN THE
ENDOTHELIUM AND SURROUNDING MATRIX AND MESENCHYME. MEDIATES BLOOD
VESSEL MATURATION/STABILITY. IT MAY PLAY AN IMPORTANT ROLE IN THE
HEART EARLY DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: Secreted.

```

```

CC -1- PTM: GLYCOSYLATED.
CC -1- MISCELLANEOUS: IT MAY HAVE A POTENTIAL THERAPEUTIC UTILITY SINCE
IT CAN BE USED FOR SPECIFICALLY TARGETING TUMOR VASCULATURE OR FOR
PROMOTING ANGIOGENIC PROCESSES IN CERTAIN ORGANS SUCH AS AN
ISCHEMIC HEART.
CC -1- SIMILARITY: CONTAINS 1 FIBRINOGEN C-TERMINAL DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use: by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; U83508; AAB50557.1; -.
DR EMBL; U83508; AAB50557.1; ALT_INIT.
DR HSSP; P02871; IFZD.
DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF00147; fibrinogen_C; 1.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00514; FIBRINAG_C_DOMAIN; 1.
KW Glycoprotein; Coiled coil; Signal; Polymorphism.
FT SIGNAL 1 15
FT CHAIN 16 498
FT DOMAIN 81 119
FT DOMAIN 153 261
FT DOMAIN 284 498
FT CARBOHYD 92 92
FT CARBOHYD 122 122
FT CARBOHYD 154 154
FT CARBOHYD 243 243
FT CARBOHYD 295 295
FT VARIANT 269 269
FT
FT
SQ SEQUENCE 498 AA; 57513 MW; 5D5FA63AEF6BE920 CRC64;
/FTID=VAR_009940.

Query Match 25.7%; Score 54.5; DB 1; Length 498;
Best Local Similarity 42.5%; Pred. No. 5.5;
Matches 17; Conservative 5; Mismatches 11; Indels 7; Gaps 2;

QY 2 NOESKLILASGGQALVNMRTYTYEK-LLWTSRVLKV 40
Db 154 NQTSRLEI-----QLLENSLSTYKLEKQLLOQTNEILKI 187

RESULT 15
AGPL_MOUSE
ID AGPL_MOUSE STANDARD; PRT; 498 AA.
AC O08538;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Angiopietin-1 precursor (ANG-1).
GN ANGPT1 OR AGPT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97134663; PubMed=8980223;
RA Davis S., Aldrich T.H., Jones P.F., Acheson A., Compton D.L., Jain V.,
RA Ryan T.E., Bruno J., Radziejewski C., Maisonnier P.C.,
RA Yancopoulos G.D.;
RT "Isolation of angiopietin-1, a ligand for the TIE2 receptor, by
secretion-trap expression cloning.";
RL Cell 87:1161-1169(1996).
CC -1- FUNCTION: BINDS AND ACTIVATES TIE2 RECEPTOR BY INDUCING ITS
TYROSINE PHOSPHORYLATION. IMPLICATED IN ENDOTHELIAL DEVELOPMENTAL

```


CC PROCESSES LATER AND DISTINCT FROM THAT OF VEGF. APPEARS TO PLAY A
CC CRUCIAL ROLE IN MEDIATING RECIPROCAL INTERACTIONS BETWEEN THE
CC ENDOTHELIUM AND SURROUNDING MATRIX AND MESENCHYME. MEDIATES BLOOD
CC VESSEL MATURATION/STABILITY. IT MAY PLAY AN IMPORTANT ROLE IN THE
CC HEART EARLY DEVELOPMENT.
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- DEVELOPMENTAL STAGE: EARLY IN DEVELOPMENT, AT E9 TO E11, IT IS
CC FOUND MOST PROMINENTLY IN THE HEART MYOCARDIUM SURROUNDING THE
CC ENDOCARDIUM. LATER, IT BECOMES MORE WIDELY DISTRIBUTED, MOST OFTEN
CC IN THE MESENCHYME SURROUNDING DEVELOPING VESSELS, IN CLOSE
CC ASSOCIATION WITH ENDOTHELIAL CELLS.
CC -|- SIMILARITY: CONTAINS 1 FIBRINOGEN C-TERMINAL DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U83509; AAB50558.1; .
DR HSP; P02671; 1F2D.
DR MGD; MGI:108448; Agpt.
DR InterPro; IPR002181; Fibrinogen_C.
DR Pfam; PF00147; fibrinogen_C; 1.
DR SMART; SM00186; FBG; 1.
DR PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; 1.
KW Glycoprotein; Coiled coil; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 498 ANGIOPOIETIN-1.
FT DOMAIN 81 119 COILED COIL (POTENTIAL).
FT DOMAIN 153 261 COILED COIL (POTENTIAL).
FT DOMAIN 284 498 FIBRINOGEN C-TERMINAL.
FT CARBOHYD 92 92 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 243 243 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 498 AA; 57505 MW; 285B4FDEC260D800 CRC64;

Query Match 25.7%; Score 54.5; DB 1; Length 498;
Best Local Similarity 42.5%; Pred. No. 5.5;
Matches 17; Conservative 5; Mismatches 11; Indels 7; Gaps 2;

Qy 2 NQESKLIILASGGPQALVNIIMRTYTYEK-LLWTTTSRLKV 40
Db 154 NQTSRLKI-----QLLENSILSTYKLEKQLLOOTNEILKI 187

Search completed: July 29, 2002, 16:09:53
Job time: 758 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 16:12:04 ; Search time 124.4 Seconds
(without alignments)
59.797 Million cell updates/sec

Title: US-09-641-104A-8
Perfect score: 212
Sequence: 1 GNQSKLIIILASGGPQALVN.....TTYEKLWTTSRVLKLSV 43

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_protein:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	212	100.0	780	13	Q90424
2	212	100.0	781	11	Q9D335
3	212	100.0	781	13	Q42486
4	207	97.6	117	6	P79321
5	197	92.9	729	13	Q9P8C8
6	195	92.0	821	5	O61229
7	194	91.5	745	4	Q15151
8	194	91.5	745	4	Q9BWC4
9	190	89.6	745	11	P70565
10	187	88.2	769	5	Q9NL44
11	187	88.2	773	5	O76152
12	174	82.1	806	5	Q25100
13	91	42.9	811	5	Q18825
14	78.5	37.0	678	5	O44326
15	57.5	27.1	364	17	Q9YDW9
16	57.5	27.1	494	3	Q9P8C8

17: 57 26.9 811 3 013705
18: 56 26.4 167 2 09X713
19: 55 25.9 842 16 098049
20: 54.5 25.7 89 11 091291
21: 54.5 25.7 170 2 069820
22: 54.5 25.7 274 11 09CST2
23: 54.5 25.7 498 6 09BDY8
24: 54 25.5 214 16 034664
25: 53 25.0 873 12 068287
26: 53 25.0 3898 12 09YRM6
27: 53 25.0 4490 4 096JB1
28: 52.5 24.8 310 5 09U1X1
29: 52 24.5 222 16 09RUP0
30: 52 24.5 312 17 0972W8
31: 52 24.5 405 17 0980P1
32: 52 24.5 721 10 092V65
33: 52 24.5 782 5 018654
34: 52 24.5 994 4 09UE60
35: 52 24.5 1928 4 099736
36: 52 24.5 2392 4 095001
37: 52 24.5 2412 4 092616
38: 52 24.5 3898 12 068534
39: 52 24.5 3898 12 068535
40: 52 24.5 3898 12 068872
41: 52 24.5 3898 12 092366
42: 52 24.5 3898 12 09Y330
43: 52 24.5 3898 12 099BK1
44: 52 24.5 3898 12 0991S8
45: 52 24.5 3898 12 068965

ALIGNMENTS

RESULT : 1

Q90424 ID Q90424 PRELIMINARY; PRT; 780 AA.
AC Q90424;
DT 01-NOV-1996 (TREMREL. 01, Created)
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)
DT 01-DEC-2001 (TREMREL. 19, Last annotation update)
DE B-CATENIN.
GN CTNNB.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96122902; PubMed=8562427;
RA Kelly,G.M., Erezylmaz D.F., Moon R.T.;
RT "Induction of a secondary embryonic axis in zebrafish occurs following the overexpression of beta-catenin."
RL Mech. Dev. 53:261-273(1995).
DR EMBL: U41081; AAC59732.1;
DR HSSP: P35222; IC3J
DR ZFIN; ZDB-GENE-980526-362; cttnb.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 12.
DR SMART; SM00185; ARM; 11.
DR PROSITE; PS0176; ARM_REPEAT; 8.
SQ SEQUENCE 780 AA; 85542 MW; D7A1FB80F94066DC CRC64;

Query Match .100.0%; Score 212; DB 13; Length 780;
Best Local Similarity 100.0%; Pred. No. 3.9e-21;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNQSKLIIILASGGPQALVNIMRTTYTEKLWTTSRVLKLSV 43
Db 306 GNQSKLIIILASGGPQALVNIMRTTYTEKLWTTSRVLKLSV 348

```

RESULT 2
Q9D335 ID Q9D335 PRELIMINARY; PRT; 781 AA.
AC Q9D335;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE ADULT MALE COLON CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,
DE CLONE:9030417H18, FULL INSERT SEQUENCE.
GN CATNB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=COLON;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikola I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK018515; BAB31250.1; -.
DR HSSP: P35222; IG3J.
DR MGD: MGI:88276; Catnb.
DR InterPro: IPR000225; Armadillo.
DR Pfam: PF00514; Armadillo_seg; 12.
DR SMART: SM00185; ARM; 12.
DR PROSITE: PS50176; ARM_REPEAT; 7.
SQ SEQUENCE 781 AA; 85546 MW; 937538C3B5CD75D1 CRC64;

Query Match 100.0%; Score 212; DB 11; Length 781;
Best Local Similarity 100.0%; Pred. No. 3.9e-21;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GQESKLIILASGGPQALVNIMRTTYEKLWTTSRVLKVLVS 43
Db 307 GQESKLIILASGGPQALVNIMRTTYEKLWTTSRVLKVLVS 349

RESULT 3
O42486 ID O42486 PRELIMINARY; PRT; 781 AA.
AC O42486;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE BETA CATENIN.
GN CHBCAT.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN-WHITE LEHORN; TISSUE=DORSAL SKIN;
RX MEDLINE=97464068; PubMed=9322759;
RA Lu J., Chuong C.M., Widelitz R.B.;
RT "Isolation and characterization of chicken beta-catenin.";
RL Gene 196:201-207(1997).
DR EMBL: U082964; AAB80856.1; -.
DR HSSP: P35222; IG3J.
DR InterPro: IPR000225; Armadillo.
DR Pfam: PF00514; Armadillo_seg; 12.
DR SMART: SM00185; ARM; 11.
DR PROSITE: PS50176; ARM_REPEAT; 8.
SQ SEQUENCE 781 AA; 85438 MW; 6D205D9A4DBAC562 CRC64;

Query Match 100.0%; Score 212; DB 13; Length 781;
Best Local Similarity 100.0%; Pred. No. 3.9e-21;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GQESKLIILASGGPQALVNIMRTTYEKLWTTSRVLKVLVS 43
Db 307 GQESKLIILASGGPQALVNIMRTTYEKLWTTSRVLKVLVS 349

RESULT 4
P79321 ID P79321 PRELIMINARY; PRT; 117 AA.
AC P79321;
DT 01-MAY-1997 (Tremblrel. 03, Created)
DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE BETA-CATENIN (FRAGMENT).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SMALL INTESTINE;
RA Winteroe A.K., Fredholm M.;
RT "Evaluation and characterization of a porcine small intestine cDNA
library.";
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z84131; CAB06327.1; -.
DR HSSP: P35222; IG3J.
DR InterPro: IPR000225; Armadillo.
DR Pfam: PF00514; Armadillo_seg; 2.
DR PROSITE: PS50176; ARM_REPEAT; 2.
FT NON_TER 1
FT NON_TER 117
SQ SEQUENCE 117 AA; 12717 MW; 73F7CCF3B917A41A CRC64;

Query Match 97.6%; Score 207; DB 6; Length 117;
Best Local Similarity 97.7%; Pred. No. 2.4e-21;
Matches 42; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GQESKLIILASGGPQALVNIMRTTYEKLWTTSRVLKVLVS 43
Db 33 GQESKLIILASGGPQALVNIMRTTYEKLWTTSRVLKVLVS 75

RESULT 5
Q9PVF7 ID Q9PVF7 PRELIMINARY; PRT; 729 AA.
AC Q9PVF7;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE CELL-ADHESION PROTEIN PLAKOGLOBIN.
GN JUP.
OS Brachydanio rerio (zebrafish) (zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 OC Cypriniformes; Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9338700; PubMed=10456847;
 RA Cerdá J., Reidenbach S., Pratzel S., Franke W.W.;
 RT "Cadherin-catenin complexes during zebrafish oogenesis: heterotypic
 RT junctions between oocytes and follicle cells.";
 RL Biol. Reprod. 61:692-704(1999).
 DR EMBL; AF099738; AAD56592.1; -;
 DR HSP; Q02248; IDOW.
 DR ZFIN; ZDB-GENE-991207-22; jup.
 DR InterPro; IPR000225; Armadillo.
 DR Pfam; PF00514; Armadillo_seg; 11.
 DR SMART; SM00185; ARM; 8.
 DR PROSITE; PS50176; ARM_REPEAT; 8.
 SQ SEQUENCE 729 AA; 80033 MW; 91E00417B4FD8CEE CRC64;

Query Match 92.9%; Score 197; DB 13; Length 729;
 Best Local Similarity 90.7%; Pred. No. 4.6e-19;
 Matches 39; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GNOESKLIILASGGPOALVNIMRTYTYEKLWTSRVLKVLVS 43
 DB 287 GNOESKLIILASGGPEGLVNIMRTYTYEKLWTSRVLKVLVS 329

RESULT 6
 OS 1 GNOESKLIILASGGPOALVNIMRTYTYEKLWTSRVLKVLVS 43
 DB 287 GNOESKLIILASGGPEGLVNIMRTYTYEKLWTSRVLKVLVS 329

OC Lytechinus variegatus (Sea urchin).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoides; Euechinoidea; Echinacea; Temnopleuroidea; Toxopneustidae;
 OC Lytechinus.
 OX NCBI_TaxID=7654;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98104237; PubMed=9441670;
 RA Miller J.R., McClay D.R.;
 RT "Changes in the pattern of adherens junction-associated beta-catenin
 RT accompany morphogenesis in the sea urchin embryo.";
 RL Dev. Biol. 192:310-322(1997).
 DR EMBL; U34814; AAC06340.1; -;
 DR HSP; Q02248; IDOW.
 DR InterPro; IPR002086; Aldehyde_dehydr.
 DR Pfam; PF00514; Armadillo_seg; 12.
 DR SMART; SM00185; ARM; 12.
 DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
 DR PROSITE; PS50176; ARM_REPEAT; 8.
 SQ SEQUENCE 821 AA; 89558 MW; 71E21D562A99C5AD CRC64;

Query Match 92.0%; Score 195; DB 5; Length 821;
 Best Local Similarity 93.0%; Pred. No. 1e-18;
 Matches 40; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GNOESKLIILASGGPOALVNIMRTYTYEKLWTSRVLKVLVS 43
 DB 324 GNOESKLIILASGGPOALVNIMRTYTYEKLWTSRVLKVLVS 366

RESULT 7
 OS 1 GNOESKLIILASGGPOALVNIMRTYTYEKLWTSRVLKVLVS 43
 DB 324 GNOESKLIILASGGPOALVNIMRTYTYEKLWTSRVLKVLVS 366

OC Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89264555; PubMed=2726765;
 RA Franke W.W., Goldschmidt M.D., Zimbelmann R., Mueller H.M.,
 RA Schiller D.L., Cowin P.;
 RT "Molecular cloning and amino acid sequence of human plakoglobin, the
 RT common junctional plaque protein.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:4027-4031(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Zimbelmann R.;
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 239-409 FROM N.A.
 RC TISSUE=PLACENTA;
 RX MEDLINE=96157724; PubMed=8576101;
 RA Ozawa M., Nuru K., Toyoyama H., Ohi Y.;
 RT "Cloning of an alternative form of plakoglobin (gamma-catenin) lacking
 RT the fourth armadillo repeat.";
 RL J. Biochem. 118:836-840(1995).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX PubMed=11016852;
 RA Whitlock N.V., Eady R.A.J., McGrath J.A.;
 RT "Genomic Organization and amplification of the human plakoglobin
 RT gene.";
 RL Exp. Dermatol. 9:323-326(2000).
 DR EMBL; Z68228; CAA92522.1; -;
 DR EMBL; D50808; BAA09435.1; -;
 DR EMBL; AF306723; AAG16727.1; -;
 DR EMBL; AF233882; AAG16727.1; JOINED.
 DR HSP; Q02248; ZBCT.
 DR InterPro; IPR000225; Armadillo.
 DR Pfam; PF00514; Armadillo_seg; 11.
 DR SMART; SM00185; ARM; 9.
 DR PROSITE; PS50176; ARM_REPEAT; 7.
 SQ SEQUENCE 745 AA; 81744 MW; 3519A0973748BCF4 CRC64;

Query Match 91.5%; Score 194; DB 4; Length 745;
 Best Local Similarity 90.7%; Pred. No. 1.3e-18;
 Matches 39; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GNOESKLIILASGGPOALVNIMRTYTYEKLWTSRVLKVLVS 43
 DB 298 GNOESKLIILASGGPOALVNIMRTYTYEKLWTSRVLKVLVS 340

RESULT 8;
 OS 1 GNOESKLIILASGGPOALVNIMRTYTYEKLWTSRVLKVLVS 43
 DB 298 GNOESKLIILASGGPOALVNIMRTYTYEKLWTSRVLKVLVS 340

OC Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LUNG CARCINOMA;
 RA Strausberg R.;

RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA, CHORTOCARCINOMA;
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC000441; AAH00441.1; -;
 DR EMBL; BC011865; AAH11865.1; -;
 DR HSSP; Q02248; 2BCT.
 DR InterPro; IPR000225; Armadillo.
 DR Pfam; PF00514; Armadillo_seg; 11.
 DR SMART; SM00185; ARM; 12.
 DR PROSITE; PS50176; ARM_REPEAT; 7.
 SQ SEQUENCE 745 AA; 81726 MW; 34DF7BFB4748BCF4 CRC64;

Query Match 91.5%; Score 194; DB 4; Length 745;
 Best Local Similarity 90.7%; Pred. No. 1.3e-18;
 Matches 39; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GNOESKLIILASGGPQALVNMRTTYEKLWTSRVKLVSV 43
 |||||:||||| ||| :|||||:|||||
 Db 298 GNOESKLIILASGGPQALVNMRTTYEKLWTSRVKLVSV 340

RESULT 9

ID P70565 PRELIMINARY; PRT; 745 AA.
 AC P70565;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE PLAKOGLOBIN.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=VAGINA;
 RA Umekita Y., Liao S.;
 RT "Molecular cloning and sequencing of the rat plakoglobin cDNA."
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=VAGINA;
 RA Hipakka R.A.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U58858; AAB06317.1; -;
 DR HSSP; Q02248; 2BCT.
 DR InterPro; IPR000225; Armadillo.
 DR Pfam; PF00514; Armadillo_seg; 11.
 DR SMART; SM00185; ARM; 8.
 DR PROSITE; PS50176; ARM_REPEAT; 7.
 SQ SEQUENCE 745 AA; 81777 MW; 9E2F52910A7ACD41 CRC64;

Query Match 89.6%; Score 190; DB 11; Length 745;
 Best Local Similarity 88.4%; Pred. No. 4.6e-18;
 Matches 38; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GNOESKLIILASGGPQALVNMRTTYEKLWTSRVKLVSV 43
 |||||:||||| ||| :|||||:|||||
 Db 298 GNOESKLIILASGGPQALVNMRTTYEKLWTSRVKLVSV 340

RESULT 10

ID Q9NL44 PRELIMINARY; PRT; 769 AA.
 AC Q9NL44;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE BETA-CATENIN.
 GN CIBETA-CATENIN.
 OS Clona; Intestinalis.
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;
 OC Clonidae; Clona.
 OX NCBI_TaxID=7719;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Imai K., Takada N., Satoh N., Satou Y.;
 RT "An essential role of beta-catenin in the endoderm specification of
 ascidian embryo."
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB031543; BAA92185.1; -;
 DR HSSP; P35222; IG3J.
 DR InterPro; IPR000225; Armadillo.
 DR Pfam; PF00514; Armadillo_seg; 11.
 DR SMART; SM00185; ARM; 11.
 DR PROSITE; PS50176; ARM_REPEAT; 6.
 SQ SEQUENCE 769 AA; 84703 MW; F61CC489B436E1BC CRC64;

Query Match 88.2%; Score 187; DB 5; Length 769;
 Best Local Similarity 88.4%; Pred. No. 1.2e-17;
 Matches 38; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GNOESKLIILASGGPQALVNMRTTYEKLWTSRVKLVSV 43
 |||||:||||| ||| :|||||:|||||
 Db 297 GNOESKLIILASGGPQALVNMRTTYEKLWTSRVKLVSV 339

RESULT 11

ID O76152 PRELIMINARY; PRT; 773 AA.
 AC O76152;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE BETA-CATENIN.
 OS Clona savignyi.
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;
 OC Clonidae; Clona.
 OX NCBI_TaxID=51511;
 [1]
 RN SEQUENCE FROM N.A.
 RC MEDLINE=98443204; PubMed=9769178;
 RA Yoshida S., Marikawa Y., Satoh N.;
 RT "Regulation of the trunk-tail patterning in the ascidian embryo: a
 RT possible interaction of cascades between lithium/beta-catenin and
 RT localized maternal factor pem."
 RL Dev. Biol. 202:264-279(1998).
 DR EMBL; AB012160; BAA32789.1; -;
 DR HSSP; P35222; IG3J.
 DR InterPro; IPR000225; Armadillo.
 DR Pfam; PF00514; Armadillo_seg; 11.
 DR SMART; SM00185; ARM; 9.
 DR PROSITE; PS50176; ARM_REPEAT; 7.
 SQ SEQUENCE 773 AA; 85217 MW; C1340CF82AFEBDAB CRC64;

Query Match 88.2%; Score 187; DB 5; Length 773;
 Best Local Similarity 88.4%; Pred. No. 1.3e-17;
 Matches 38; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GNOESKLIILASGGPQALVNMRTTYEKLWTSRVKLVSV 43
 |||||:||||| ||| :|||||:|||||
 Db 296 GNOESKLIILASGGPQALVNMRTTYEKLWTSRVKLVSV 338

RESULT 12

ID Q25100 PRELIMINARY; PRT; 806 AA.
 AC Q25100;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE BETA-CATENIN.
 OS Hydra magnipapillata (Hydra).
 OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;
 OC Hydridae; Hydra.
 OX NCBI_TaxID=6085;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=WILDTYPE 105;
 RX MEDLINE=96257271; PubMed=8654597;
 RA Hombayer E., Hatt M., Fischer R., Fujisawa T., Holstein T.W.,
 RA Sugiyama T.;
 RT "Identification of a Hydra homologue of the beta-
 catenin/plakoglobin/armadillo gene family.";
 RL Gene 172:155-159(1996).
 DR EMBL: U36781; AAC47137.1; -.
 DR HSSP: Q02248; 2BCT.
 DR InterPro: IPR000225; Armadillo.
 DR Pfam: PF00514; Armadillo_seg; 10.
 DR SMART: SM00185; ARM; 10.
 DR PROSITE: PS50176; ARM_REPEAT; 7.
 SQ SEQUENCE 806 AA; 90462 MW; 689E5E982CD5051A CRC64;

Query Match 82.1%; Score 174; DB 5; Length 806;
 Best Local Similarity 83.7%; Pred. No. 8.8e-16;
 Matches 36; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 GNQESKLIILASGGPQALVNMIRTYTYEKLWTTSRVLKLSV 43
 |||||
 DB 366 GNQESKLIILSSGGPQALVNMIRSYTYEKLWTTTCRVLKLSV 408

RESULT 13

ID Q18825 PRELIMINARY; PRT; 811 AA.
 AC Q18825;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE C54D1.6 PROTEIN.
 GN C54D1.6 OR BAR-1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
 RA Bonfield J., Burton J., Connell M., Copsey J., Cooper J., Fulton L.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
 RA Smaildon N., Smith A., Sonnenhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans.";
 RL Nature 368:32-38(1994).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Minx M.;
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=N2;
 RA Eisenmann D.M., Malloof J.N., Simske J.S., Kenyon C., Kim S.K.;

RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U46673; AAC48154.1; -.
 DR EMBL: AF063646; AAC17424.1; -.
 DR HSSP: Q02248; 3BCT.
 DR TRANSFAC: T03880; -.
 SQ SEQUENCE 811 AA; 92227 MW; 97D6FFDE71BDFFDFF CRC64;

Query Match 42.9%; Score 91; DB 5; Length 811;
 Best Local Similarity 41.5%; Pred. No. 0.0004;
 Matches 17; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 2 NOESKLIILASGGPQALVNMIRTYTYEKLWTTSRVLKLSV 42
 |||||
 DB 290 NTEOKIKFKVKGSGPQKLLMLQHRVVENLLWRTTQLKTF 330

RESULT 14

ID O44326 PRELIMINARY; PRT; 678 AA.
 AC O44326;
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE HMP-2.
 GN HMP-2 OR K05C4.6.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=N2 BRISTOL;
 RA Costa M., Raich W., Agunag C., Hardin J., Priess J.R.;
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Harris B.;

RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF016853; AAB94552.1; -.
 DR EMBL: Z81584; CAB04572.1; -.
 DR HSSP: P35222; IG3J.
 DR InterPro: IPR000225; Armadillo.
 DR Pfam: PF00514; Armadillo_seg; 6.
 DR PROSITE: PS50176; ARM_REPEAT; 1.
 SQ SEQUENCE 678 AA; 74510 MW; E6C7ED51F6241232 CRC64;

Query Match 37.0%; Score 78.5; DB 5; Length 678;
 Best Local Similarity 45.0%; Pred. No. 0.019;
 Matches 18; Conservative 10; Mismatches 11; Indels 1; Gaps 1;

QY 5 SKLIILASGGPQALVNMIRTYTYEKLWTTSRVLKLSV 43
 |||||
 DB 239 SKITFLLSGPQILVSLVSDHRKLIIVTVRCIRSLV 278

RESULT 15

Q9YDW9 PRELIMINARY; PRT; 364 AA.
 ID Q9YDW9;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE HYPOTHETICAL 38.4 KDA PROTEIN APE0800.
 GN APE0800.
 OS Aeropyrum pernix.
 OC Archaea; Crenarchaeota; Desulfurococcaceae; Desulfurococcaceae;
 OC Aeropyrum.
 OX NCBI_TaxID=56636;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K1;

Query Match	27.1%;	Score 57.5;	DB 17;	Length 364;
Best Local Similarity	39.5%;	pred. No. 8.2;		
Matches 15;	Conservative 7;	Mismatches 9;	Indels 7;	Gaps 1;

QY

1 GNQESKLIILASG-----GPQAIVNIMRTTYTEKL 31
| | : | | : | |
| | : | | : | | :

db

173 GSREFDAIIAAGPWTGTISGLQALSNMLRVYRCEALI 210

Search completed: July 29, 2002, 16:12:06
Job time: 476 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 16:07:54 ; Search time 158.47 seconds

(without alignments)

28.737 Million cell updates/sec

Title: US-09-641-104A-9

Perfect score: 217

Sequence: 1 CSSNKPATVEAGMGQALGLH.....TDPQRLVQNCWTLRNLSL 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

1: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	DB ID	Description
1	217	100.0	41	20 AAY33220 Human beta-catenin
2	217	100.0	41	20 AAY33235 Human beta-catenin
3	217	100.0	781	21 AAB07290 Human beta-catenin
4	217	100.0	781	21 AAY70740 Human beta-catenin
5	217	100.0	781	22 AAU28118 Novel human secret
6	217	100.0	781	22 AAE06038 Human beta-catenin
7	217	100.0	781	22 AAE06039 Mouse beta-catenin
8	217	100.0	800	22 AAU28306 Novel human secret
9	185	85.3	700	12 AAR11354 Cellular Receptor
10	179	82.5	840	22 ABB60196 Drosophila melanog
11	179	82.5	840	22 ABB65819 Drosophila melanog

12	179	82.5	840	22 ABB65821 Drosophila melanog
13	72.5	33.4	799	21 AAY92061 Human APC-2 polype
14	67.5	31.1	2274	22 AAB50674 Mouse APC-2 protei
15	61	28.1	542	22 AAB70385 Yeast host cell pr
16	60	27.6	522	22 ABB60982 Drosophila melanog
17	59.5	27.4	45	22 ABB38201 Peptide #5707 enco
18	59.5	27.4	45	22 ABB23386 Protein #5385 enco
19	59.5	27.4	45	22 AAM58834 Human brain expres
20	59.5	27.4	45	22 AAM71352 Human bone marrow
21	59.5	27.4	45	22 AAM13004 Peptide #5438 enco
22	59.5	27.4	45	22 AAM31632 Peptide #5669 enco
23	59.5	27.4	584	18 AAW09647 Mouse merlin prote
24	59.5	27.4	584	20 AAW94457 Mouse neurofibroma
25	59.5	27.4	590	20 AAW94458 Human neurofibroma
26	59.5	27.4	591	18 AAW09646 Mouse merlin prote
27	59.5	27.4	591	18 AAW09648 Human merlin prote
28	59.5	27.4	591	20 AAW94456 Mouse neurofibroma
29	59.5	27.4	595	15 AAW60398 Merlin protein enc
30	59.5	27.4	596	18 AAW09645 Mouse merlin prote
31	59.5	27.4	596	20 AAW94455 Mouse neurofibroma
32	59.5	27.4	596	20 AAW94459 Human neurofibroma
33	57.5	26.5	539	22 ABB64913 Drosophila melanog
34	57	26.3	767	19 AAW46272 Moraxella catarrha
35	56	25.8	779	22 ABB66318 Drosophila melanog
36	55	25.3	296	22 AAU38312 Salmonella typhi c
37	55	25.3	821	21 AAB36467 Xenopus laevis pla
38	54.5	25.1	323	20 AAY21811 CPF polypeptide 36
39	54	24.9	137	22 AAG18834 Novel human diagno
40	54	24.9	215	22 AAG18836 Novel human diagno
41	54	24.9	318	22 AAG81387 Human AFP protein
42	54	24.9	330	22 AAG81379 Human AFP protein
43	54	24.9	353	22 ABB70617 Drosophila melanog
44	54	24.9	453	18 AAW16328 Human host cell pr
45	54	24.9	453	22 AAB70386 Human host cell pr

ALIGNMENTS

RESULT 1:

AY33220
ID AAY33220 standard; peptide; 41 AA.

XX AC AAY33220;

XX DT 18-NOV-1999 (first entry)

XX DE Human beta-catenin protein armadillo repeat arm6 fragment.

XX KW Beta-catenin; human; armadillo repeat; treatment; human disease; LEF-1; modulator; interaction domain; transcription factor; TCF-4; oncogenic;
KW tumor suppressor gene; conductin; E-cadherin; screening; carcinoma; APC;
KW Wnt signalling pathway; tumor development; anti-oncogenic; melanoma;
XX organ regeneration; tissue regeneration; hair growth.

XX OS Homo sapiens.

XX PN DE19909251-Al.

XX PD 26-AUG-1999.

XX PF 22-FEB-1999; 99DE-1009251.

XX PR 21-FEB-1998; 98DE-1007390.

XX PA (DELB-) DELBRUECK CENT.MOLEKULARE MEDIZIN MAX.

XX PI Birchmeier W, Von Kries J;

XX WPI; 1999-470389/40.

XX DR Agents for treating human diseases, particularly cancer, modulate

XX PT interaction of beta-catenin with transcription factors or tumor

PT suppressor gene products
XX
PS Disclosure; Page 7; 16pp; German.
XX
XX This invention describes a novel agent (A) for treating human disease
CC which is based on substances (I) that modulate (inhibit or promote) the
CC interaction of beta-catenin with transcription factors or products of
CC tumor suppressor genes. The invention also describes (a) peptides (II)
CC comprising part of the LEF-1/TCF-4 transcription factors, or their
CC variants and mutants; (b) peptides and related molecules (III) from the
CC armadillo domain (arm units 3-8) of beta-catenin, and mutants of the
CC complete beta-catenin molecule, that include at least one of the
CC specific interaction domains for LEF-1, TCF-4, APC, conductin or
CC E-cadherin; and (c) enzyme-linked immunosorbent assay (ELISA) for
CC screening substance libraries for compounds that modulate interaction of
CC beta-catenin with LEF-1/TCF, APC, conductin or E-cadherin. Beta-catenin
CC is a key compound in the Wnt signalling pathway and is involved in
CC development of tumors. Generally its interaction with LEF-1 or TCF-4 is
CC oncogenic but interaction with APC, conductin or E-cadherin is
CC anti-oncogenic. (A) which inhibit interaction are particularly used to
CC treat tumors, especially carcinoma of the colon and melanoma, but also,
CC where they promote interaction, to stimulate regeneration of organs and
CC tissues, specifically hair growth. AAY3230-Y3241 represent mutant
CC human beta-catenin armadillo repeat fragments described in the method of
CC the invention.
XX
SQ Sequence 41 AA;
Query Match 100.0%; Score 217; DB 20; Length 41;
Best Local Similarity 100.0%; Pred. No. 9.3e-25;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CSSNKPATVEAGGMOALGLHLTPDSORLVQNCWLTLRLNSD 41
Db 1 cssnkpaveagmgalghltopsgqlvncqlwtlrlnsd 41
RESULT 3
AAB07290
ID AAB07290 standard; Protein; 781 AA.
XX
AC AAB07290;
XX
XX 01-NOV-2000 (first entry)
XX Human beta catenin.
XX Beta catenin; cadherin; metastasis; cadherin-associated protein;
KW human; colorectal cancer; melanoma; antisense oligonucleotide;
KW gene therapy.
XX Homo sapiens.
XX OS US6066500-A.
XX PN 23-MAY-2000.
XX PD 25-JUN-1999; 99US-0344519.
XX PF 25-JUN-1999; 99US-0344519.
XX PR (ISIS-) ISIS PHARM INC.
XX PA Bennett CF, Cowsett LM;
XX PI WPI; 2000-410651/35..
XX DR N-PSDB; AAA58320.
XX PT New antisense compounds targeting nucleic acids encoding human beta
XX catenin (HBC) useful for treating diseases associated with HBC
XX expression and as prophylaxis to prevent or delay infection,
XX inflammation or tumor formation
XX PT

PT suppressor gene products
XX
PS Disclosure; Page 7; 16pp; German.
XX
XX This invention describes a novel agent (A) for treating human disease
CC which is based on substances (I) that modulate (inhibit or promote) the
CC interaction of beta-catenin with transcription factors or products of
CC tumor suppressor genes. The invention also describes (a) peptides (II)
CC comprising part of the LEF-1/TCF-4 transcription factors, or their
CC variants and mutants; (b) peptides and related molecules (III) from the
CC armadillo domain (arm units 3-8) of beta-catenin, and mutants of the
CC complete beta-catenin molecule, that include at least one of the
CC specific interaction domains for LEF-1, TCF-4, APC, conductin or
CC E-cadherin; and (c) enzyme-linked immunosorbent assay (ELISA) for
CC screening substance libraries for compounds that modulate interaction of
CC beta-catenin with LEF-1/TCF, APC, conductin or E-cadherin. Beta-catenin
CC is a key compound in the Wnt signalling pathway and is involved in
CC development of tumors. Generally its interaction with LEF-1 or TCF-4 is
CC oncogenic but interaction with APC, conductin or E-cadherin is
CC anti-oncogenic. (A) which inhibit interaction are particularly used to
CC treat tumors, especially carcinoma of the colon and melanoma, but also,
CC where they promote interaction, to stimulate regeneration of organs and
CC tissues, specifically hair growth. AAY3217-Y3222 represent human
CC beta-catenin armadillo repeat fragments described in the method of the
CC invention.
XX
SQ Sequence 41 AA;
Query Match 100.0%; Score 217; DB 20; Length 41;
Best Local Similarity 100.0%; Pred. No. 9.3e-25;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CSSNKPATVEAGGMOALGLHLTPDSORLVQNCWLTLRLNSD 41
Db 1 cssnkpaveagmgalghltopsgqlvncqlwtlrlnsd 41
RESULT 2
AAY32325
ID AAY32325 standard; peptide; 41 AA.
XX
AC AAY32325;
XX
XX 18-NOV-1999 (first entry)
XX Human beta-catenin protein mutant armadillo repeat arm 6.
XX Beta-catenin; human; armadillo repeat; treatment; human disease; LEF-1;
KW modulator; interaction domain; transcription factor; TCF-4; oncogenic;
KW tumor suppressor gene; conductin; E-cadherin; screening; carcinoma; APC;
KW Wnt signalling pathway; tumor development; anti-oncogenic; melanoma;
KW organ regeneration; tissue regeneration; hair growth; mutant.
XX
XX Homo sapiens.
XX OS Synthetic.
XX PN DE19909251-A1.
XX XX 26-AUG-1999.
XX PD 22-FEB-1999; 99DE-1009251.
XX PF 21-FEB-1998; 98DE-1007390.
XX PR (DELB-) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.
XX PA Birchmeier W, Von Kries J;
XX PI WPI; 1999-470389/40.
XX DR Agents for treating human diseases, particularly cancer, modulate
XX interaction of beta-catenin with transcription factors or tumor
XX PT

XX Example 13; Columns 45-52; 35pp; English.

XX Beta catenin is a member of the catenin family of cytosolic proteins and

CC a key member of the Wnt signalling pathway. Catenins interact with the

CC cytoplasmic domains of cadherin glycoproteins, and are important in

CC maintaining cell adhesiveness. The loss of cell adhesiveness is

CC implicated in metastasis. Beta catenin is also known as

CC cadherin-associated protein and is implicated in colorectal cancer and

CC melanoma. The present sequence is the human beta catenin protein. The

CC coding sequence of this protein was used in the present invention to

CC design antisense oligonucleotides (AA58327-A58366). The

CC oligonucleotides are capable of hybridising to human beta catenin, in

CC order to inhibit expression of human beta catenin. The oligonucleotides

CC may be used in gene therapy for colorectal cancer or melanoma.

XX Sequence 781 AA;

SQ

Query Match 100.0%; Score 217; DB 21; Length 781;

Best Local Similarity 100.0%; Pred. No. 3.6e-23;

Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSSNKPAAIVEAGGMOALGHLTDPQSRLVQNCWLTNRNLS 41

Db 350 cssnkpaaiveaggmqalghltdpsqrlvqncwltlnl 390

RESULT 5

AAU28118

ID AAU28118 standard; Protein; 781 AA.

XX

AC AAU28118;

XX

DT 18-DEC-2001 (first entry)

XX

DE Novel human secretory protein, Seq ID No 287.

XX

KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;

KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;

KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;

KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;

KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;

KW gut protection; lung; liver fibrosis; immune deficiency; infection;

KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;

KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;

KW fertility; analgesic; pain; antigen.

XX

OS Homo sapiens.

XX

PN WO200166689-A2.

XX

PD 13-SEP-2001.

XX

PF 05-MAR-2001; 2001WO-US04942.

XX

PR 07-MAR-2000; 2000US-0519705.

PR 19-MAY-2000; 2000US-0574454.

PR 17-JUN-2000; 2000US-0596193.

PR 14-JUL-2000; 2000US-0616847.

PR 19-SEP-2000; 2000US-0665363.

PR 20-OCT-2000; 2000US-0693267.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;

PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;

XX

XX WPI: 2001-589934/66.

DR N-PSDS; AAS45018.

XX

PT Novel-polypeptides and nucleic acids obtained from cDNA libraries

PT prepared from various human tissues, for diagnosis and treatment of

PT cancer, neurological, inflammatory, and autoimmune disorders -

XX

PS Example 4; SEQ ID No 287; 107pp; English.

XX The invention relates to novel isolated human secreted polypeptides (I)

CC and polynucleotides (II). (I) and (II) are useful for treating

CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,

CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is

CC involved in increasing haematopoiesis, stem cell survival, bone growth

CC and remodeling. (I), (II) and modulators of (II) are useful for

CC prophylaxis or treatment of one or more cancers. (II) is also useful for

CC creating transgenic animals useful for studying the in vivo activities of

the polypeptide as well as for studying modulators of the polypeptides.
 (1) induces the proliferation of neural cells and regeneration of nerve
 and brain tissue and is useful for the treatment of central and
 peripheral nervous system diseases and neuropathies, such as Alzheimer's,
 Parkinson's disease, Huntington's disease, and amyotrophic lateral
 sclerosis. In addition, (1) is involved in chemotactic or chemokinetic
 activity, regulation of haematopoiesis and is useful for treating myeloid
 or lymphoid cell disorders, platelet disorders such as thrombocytopenia
 and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 tissue growth, and in tissue repair, healing of burns, incisions,
 ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
 disorders, or periodontal disease. Furthermore, (1) is also useful for
 gut protection or regeneration and treatment of lung or liver fibrosis,
 reperfusion injury in various tissues, various immune deficiencies and
 disorders including severe combined immunodeficiency (SCID), bacterial or
 fungal infections, autoimmune disorders e.g. multiple sclerosis,
 rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
 reactions and conditions, such as asthma or other respiratory problems.
 In addition, (1) affects biorhythms or circadian cycles of rhythms,
 fertility, metabolism, catabolism, anabolism, storage or elimination of
 dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
 analgesic effects or other pain reducing effects, immunoglobulin like
 activity and can act as an antigen in a vaccine composition to raise an
 immune response. AAU28020-AAU28395 represent novel human secreted protein
 amino acid sequences of the invention.

XX Sequence 781 AA;

Query Match 100.0%; Score 217; DB 22; Length 781;
 Best Local Similarity 100.0%; Pred. No. 3.6e-23;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSSNKPAAIVEAGGQALGLHLDPSQRLVQNCIWLTLRLNSD 41
 DB 350 cssnkpaaiveaggmqalghldtspqrlvqnciwltrlnsd 390

RESULT 6

AAE06038

ID AAE06038 standard; Protein; 781 AA.

AC AAE06038;

DT 25-SEP-2001 (first entry)

DE Human beta-catenin protein #1.

Human; stem cell culturing; progenitor cell; pluripotential phenotype;
 transplantation; haematopoietic function; allogeneic recipient;
 signalling pathway; beta-catenin.

OS Homo sapiens.

XX WO200152649-A1.

PN 26-JUL-2001.

PF 17-JAN-2001; 2001WO-US01459.

PR 18-JAN-2000; 2000US-0176786.

XX (STRD) UNIV LELAND STANFORD JUNIOR.

PI Reya T, Nusse R, Weissman IL;

XX WPI; 2001-465328/50.

DR N-PSDB; AAD11164.

In vitro expansion of mammalian stem or progenitor cells, useful for
 producing cells that retain their pluripotential phenotype after
 expansion for use in transplantation, involves increasing the levels of
 beta-catenin in the cell -

XX Disclosure; Page 23-24; 33pp; English.
 XX The present invention relates to a method for in vitro expansion of
 CC mammalian stem or progenitor cells, which comprising increasing the
 CC intracellular concentration of beta-catenin in a progenitor or stem cell
 CC in an in vitro culture medium for a period sufficient for the progenitor
 CC or stem cell to divide. The number of cells having the functional
 CC phenotype of the stem or progenitor cells is expanded. The method is
 CC useful for culturing stem cells and progenitor cells, which retain their
 CC pluripotential phenotype after expansion, in vitro. The expanded cell
 CC populations are useful as a source of stem cells, e.g. to reconstitute
 CC function in a host that is deficient in a particular cell lineage or
 CC lineages. The expanded cell populations are also useful in
 CC transplantation to restore haematopoietic function to autologous or
 CC allogeneic recipients. The present sequence is human beta-catenin
 CC protein. Beta-catenin is a pivotal player in the signalling pathway
 CC initiated by Wnt proteins, which are mediators of several developmental
 CC processes.
 XX Sequence 781 AA;

Query Match 100.0%; Score 217; DB 22; Length 781;

Best Local Similarity 100.0%; Pred. No. 3.6e-23;

Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSSNKPAAIVEAGGQALGLHLDPSQRLVQNCIWLTLRLNSD 41

DB 350 cssnkpaaiveaggmqalghldtspqrlvqnciwltrlnsd 390

RESULT 7

AAE06039

ID AAE06039 standard; Protein; 781 AA.

AC AAE06039;

DT 25-SEP-2001 (first entry)

DE Mouse beta-catenin protein #2.

Mouse; stem cell culturing; progenitor cell; pluripotential phenotype;
 transplantation; haematopoietic function; allogeneic recipient;
 signalling pathway; beta-catenin.

OS Mus musculus.

XX WO200152649-A1.

PN 26-JUL-2001.

PF 17-JAN-2001; 2001WO-US01459.

PR 18-JAN-2000; 2000US-0176786.

XX (STRD.) UNIV LELAND STANFORD JUNIOR.

PI Reya T, Nusse R, Weissman IL;

XX WPI; 2001-465328/50.

DR N-PSDB; AAD11165.

In vitro expansion of mammalian stem or progenitor cells, useful for
 producing cells that retain their pluripotential phenotype after
 expansion for use in transplantation, involves increasing the levels of
 beta-catenin in the cell -

PS Disclosure; Page 28-30; 33pp; English.

XX The present invention relates to a method for in vitro expansion of
 CC mammalian stem or progenitor cells, which comprising increasing the
 CC intracellular concentration of beta-catenin in a progenitor or stem cell

in an in vitro culture medium for a period sufficient for the progenitor or stem cell to divide. The number of cells having the functional phenotype of the stem or progenitor cells is expanded. The method is useful for culturing stem cells and progenitor cells, which retain their pluripotential phenotype after expansion, in vitro. The expanded cell populations are useful as a source of stem cells, e.g. to reconstitute function in a host that is deficient in a particular cell lineage or lineages. The expanded cell populations are also useful in transplantation to restore haematopoietic function to autologous or allogeneic recipients. The present sequence is mouse beta-catenin protein. Beta-catenin is a pivotal player in the signalling pathway initiated by Wnt proteins, which are mediators of several developmental processes.

XX Sequence 781 AA;

Query Match 100.0%; Score 217; DB 22; Length 781;
Best Local Similarity 100.0%; Pred. No. 3.6e-23;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CSSNKPALVEAGGMOALGLHLTDPQSRLVQNCIWLRLNSD 41

Db 350 cssnkpaveagggmqalghltdpsqrlvqnciwlrlnsd 390

RESULT 8

AAU283106
ID AAU28306 standard; Protein; 800 AA.

XX AAU28306;

XX 18-DEC-2001 (first entry)

XX Novel human secretory protein, Seq ID No 663.

XX Human; secreted protein; arthritis; Crohn's disease; sepsis; shock; ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy; transgenic animal; Alzheimer's disease; Parkinson's disease; burn; amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia; ulcer; osteoporosis; bone degenerative disorder; periodontal disease; gut protection; lung; liver fibrosis; immune deficiency; infection; severe combined immunodeficiency; SCID; autoimmune disorder; allergy; multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma; fertility; analgesic; pain; antigen.

XX Homo sapiens.

XX WO200166689-A2.

XX 13-SEP-2001.

XX 05-MAR-2001; 2001WO-US04942.

XX 07-MAR-2000; 2000US-0519705.

XX 19-MAY-2000; 2000US-0574454.

XX 17-JUN-2000; 2000US-0596193.

XX 14-JUL-2000; 2000US-0616847.

XX 19-SEP-2000; 2000US-0665363.

XX 20-OCT-2000; 2000US-0693267.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;

XX WPI; 2001-589934/66.

XX N-PSDB; AA545206.

XX Novel polypeptides and nucleic acids obtained from cDNA libraries
PT prepared from various human tissues, for diagnosis and treatment of
PT cancer, neurological, inflammatory, and autoimmune disorders -

PS Example 2; SEQ ID No 663; 107pp; English.

XX The invention relates to novel isolated human secreted polypeptides (I)
CC and polynucleotides (II). (I) and (II) are useful for treating
CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
CC involved in increasing haematopoiesis, stem cell survival, bone growth
CC and remodeling. (I); (II) and modulators of (II) are useful for
CC prophylaxis or treatment of one or more cancers. (II) is also useful for
CC creating transgenic animals useful for studying the in vivo activities of
CC the polypeptide as well as for studying modulators of the polypeptides.
CC (I) induces the proliferation of neural cells and regeneration of nerve
CC and brain tissue and is useful for the treatment of central and
CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
CC activity, regulation of haematopoiesis and is useful for treating myeloid
CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
CC tissue growth, and in tissue repair, healing of burns, incisions,
CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
CC disorders, or periodontal disease. Furthermore, (I) is also useful for
CC gut protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues, various immune deficiencies and
CC disorders including severe combined immunodeficiency (SCID), bacterial or
CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
CC reactions and conditions, such as asthma or other respiratory problems.
CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
CC fertility, metabolism, catabolism, anabolism, storage or elimination of
CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
CC analgesic effects or other pain reducing effects, immunoglobulin like
CC activity and can act as an antigen in a vaccine composition to raise an
CC immune response. AAU28020-AAU28395 represent novel human secreted protein
CC amino-acid sequences of the invention.

XX Sequence 800 AA;

Query Match 100.0%; Score 217; DB 22; Length 800;

Best Local Similarity 100.0%; Pred. No. 3.7e-23;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CSSNKPALVEAGGMOALGLHLTDPQSRLVQNCIWLRLNSD 41

Db 363 cssnkpaveagggmqalghltdpsqrlvqnciwlrlnsd 403

RESULT 9

AAU11354

ID AAU11354 standard; Protein; 700 AA.

XX AAU11354;

XX 03-JUN-1991 (first entry)

XX Cellular Receptor 2 ligand.

DE Cellular receptor 2; CR2; binding site; BS; auto-immune disease;
KW Epstein-Barr Virus; EBV; B lymphocyte; ligand.

XX Synthetic.

XX Key Location/Qualifiers

FT Binding-site 300..312

FT /note= "fragment pref. included"

FT Binding-site 303..309

FT /note= "fragment most pref. included"

FT Binding-site 385..400

FT /note= "fragment pref. included"

FT Binding-site 389..394

FT /note= "fragment most pref. included"

XX

CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB116176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 840 AA;

Query Match 82.5%; Score 179; DB 22; Length 840;
 Best Local Similarity 82.9%; Pred. No. 1.7e-17;
 Matches 34; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CSSNKPATVEAGGMOALGLHLTDPQSRLVQNCWLTRLNLS 41
 |||||:|||||:|:|:|:|||||:|||||
 Db 355 cssnkpavdaggmqalamhlgmnsprlvqncwltrlnsd 395

RESULT 12

ABB65821
 ID ABB65821 standard; Protein; 840 AA.

XX AC ABB65821;

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 24255.

XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR N-PSDB; ABL09924.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

XX Disclosure; SEQ ID NO 24255; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).

XX The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 840 AA;

Query Match 82.5%; Score 179; DB 22; Length 840;
 Best Local Similarity 82.9%; Pred. No. 1.7e-17;
 Matches 34; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CSSNKPATVEAGGMOALGLHLTDPQSRLVQNCWLTRLNLS 41
 |||||:|||||:|:|:|:|||||:|||||
 Db 355 cssnkpavdaggmqalamhlgmnsprlvqncwltrlnsd 395

RESULT 13

AAAY2061
 ID AAAY2061 standard; Protein; 799 AA.

XX AC AAAY2061;

DT 01-AUG-2000 (first entry)

DE Human APC-2 polypeptide.

XX Adenomatous polyposis coli 2; APC-2; tumor suppressor; Wnt pathway;
 KW cytosolic.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 60 /note= "encoded by stop codon"

XX WO200018913-A1.

XX 06-APR-2000.

XX 24-SEP-1999; 99WO-NL00595.

XX 25-SEP-1998; 98EP-0203237.

XX (UYUT-) RIJKSUNIV UTRECHT.

XX Van Es JH, Pelfer MA, Clevers JC;

XX WPI; 2000-293152/25.

DR N-PSDB; AAA08835.

XX Novel human tumor suppressor gene, adenomatous polyposis coli (APC)-2,
 PT used in the diagnosis, prevention, and treatment of cancer

XX Claim 8; Fig 4; 43pp; English.

XX This is the human adenomatous polyposis coli 2 (APC-2) polypeptide. The
 CC human APC-2 gene (AAA08835) is located at chromosomal position 19p13.3.
 CC The APC-2 tumor suppressor gene controls the Wnt pathway by the
 CC formation of a complex with GSK-3b, axin/conductin and beta-catenin,
 CC inducing the rapid degradation of the latter. The APC-2 polynucleotides
 CC are used for the diagnosis (claimed), prevention and treatment of cancer.
 CC They are also used for the recombinant production of APC-2 proteins.
 CC The methods are used to identify candidate drugs (claimed), especially
 CC for use in the treatment of cancers, preferably a fetal cancer, cancers
 CC with a stem cell phenotype, brain, lung, kidney or intestinal cancer (all
 CC claimed).

XX SQ Sequence 799 AA;

Query Match 33.4%; Score 72.5; DB 21; Length 799;
 Best Local Similarity 32.3%; Pred. No. 0.091;
 Matches 21; Conservative 4; Mismatches 15; Indels 25; Gaps 1;

Qy 1 CSSNKPATVEAGG-----MQALGLHLTDPQSRLVQNCWLMT 35
 |||||:|||||:|:|:|:|||||:|||||
 Db 656 cgsnslaliesgggilrnsvsilvatretyrqlrdhncqltqlghtshltivsnacgt 715

QY 36 LRNLS 40
| | | |
Db 716 lwnls 720

RESULT 14

AAB50674
ID AAB50674 standard; Protein: 2274 AA.

AC AAB50674;

DT 19-MAR-2001 (first entry)

DE Mouse APC-2 protein sequence SEQ ID NO:65.

KW Human; Caenorhabditis elegans; UNC-5; splice variant; nematode worm;
protein-protein interaction; identification.

OS Mus sp.

PN WO200073328-A2.

PD 07-DEC-2000.

PF 02-JUN-2000; 2000WO-EP05108.

PR 01-JUN-1999; 99GB-0012755.

PA (DEVG-) DEVGEN NV.

PI Van Criekeinge W, Roelens I, Bogaert T, Verwaerde P;

PP WPI; 2001-016508/02.

PT Three variants of human unc-5C cDNAs (unc-5Cb, unc-5Cc and unc-5C8) and
a human unc-5H51 cDNA, useful in yeast two hybrid experiments for
identifying unknown human cDNAs which encode proteins that interact
with the human unc-5C protein.

PS Disclosure; Page 185-192; 246pp; English.

CC The present invention describes 3 variants of human unc-5C cDNAs
(unc-5Cb, unc-5Cc and unc-5C8) which correspond to alternatively spliced
unc-5C transcripts, and a human unc-5H51 cDNA which shares homology with
the Rattus norvegicus unc-5H51 cDNA. Also described are assays based on
protein-protein interactions between the unc-5 protein and a variety of
different interacting proteins. The unc-5C variant cDNAs and unc-5H51
cDNA are useful in methods for identifying compounds which reduce or
inhibit the lethal phenotype associated with the expression of the
unc-5 death domain in yeast. They are also useful in yeast two hybrid
experiments for identifying unknown human cDNAs which encode proteins
that interact with the human unc-5C protein. AAC90914 to AAC90971 and
AAB50646 to AAB50693 represent sequences used in the exemplification of
the present invention.

XX Sequence 2274 AA;

QY Query Match 31.18; Score 67.5; DB 22; Length 2274;

Best Local Similarity 29.2%; Pred. No. 1.8;

Matches 19; Conservative 5; Mismatches 16; Indels 25; Gaps 1;

QY 1 CSSNKPATVEAGG-----MQALGLHLTDPSQRLVQNCILWT 35

Db 582 cgnslaviesgggllrnvssliatredyrgvldhncqlqlqltshslitvsnacgt 641

QY 36 LRNLS 40

| | | |

Db 642 lwnls 646

RESULT 15

AAB70385

ID AAB70385 standard; Protein: 542 AA.

AC AAB70385;

DT 02-MAY-2001 (first entry)

DE Yeast host cell protein SRP1 SEQ ID NO:5.

KW Identification; antiviral; viral protein; viral replication; NP;

KW viral infection; nucleoprotein.

OS Saccharomyces cerevisiae.

PN WO200111335-A2.

PD 15-FEB-2001.

PF 11-AUG-2000; 2000WO-US22257.

PR 11-AUG-1999; 99US-0148263.

XX (MOUN) MOUNT SINAI SCHOOL MEDICINE.

XX O'Neill R, Harty R, Palese PM;

XX WPI; 2001-168816/17.

XX Identifying a substance that inhibits the interaction between a viral
protein and a host cell protein, useful for the discovery of new
antiviral compounds.

XX Example; Fig 3; 147pp; English.

CC The present invention describes a method (M1) for identifying a
substance that inhibits the interaction of a viral protein (VP) with a
host cell protein (HP). The method comprises: (a) contacting HP with VP
in the presence of a test substance; and (b) detecting complex formation,
where the ability of the test substance to inhibit HP/VP interaction is
indicated by a decrease in complex formation. The antiviral compounds
that inhibit the interaction between a host protein (NSI-HP or NPI-1)
and a viral protein (NSI) are useful for treating or inhibiting viral
infection, preferably influenza and rhabdovirus infection, in humans.
Antiviral compounds include peptides and antibodies. In particular
compositions comprising a polypeptide containing an amino acid sequence
corresponding to the NP-NLS domain of the influenza virus NP protein,
which inhibits the specific interaction of the NPI-1 protein with the
influenza virus NP protein are useful for treating or inhibiting
influenza viral infection in humans. The present sequence represents
a yeast host cell protein designated SRP1, which is used in an example
from the present invention.

XX Sequence 542 AA.

QY Query Match 28.1%; Score 61; DB 22; Length 542;

Best Local Similarity 30.3%; Pred. No. 2.8;

Matches 10; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

QY 8 IVEAGGMOALGLHLTDPSQRLVQNCILWTLRNLS 40

Db 337vinagvlpalrlllsspkknkacwtisnit 369

Search completed: July 29, 2002, 16:07:55

Job time: 645 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 16:05:09 ; Search time 57.6 seconds
(without alignments)
17.386 Million cell updates/sec

Title: US-09-641-104A-9

Perfect score: 217

Sequence: 1 CSSNKPAIVEAGGMOALGLH.....TDPQRLVQNCWLTLRLNSD 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*

1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/PCTUS.COMB.pep.*

6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	28.1	542	1	US-08-246-583-3
2	59.5	27.4	584	1	US-08-179-738-7
3	59.5	27.4	584	2	US-08-628-145-7
4	59.5	27.4	591	1	US-08-179-738-5
5	59.5	27.4	591	1	US-08-179-738-10
6	59.5	27.4	591	2	US-08-628-145-5
7	59.5	27.4	591	2	US-08-628-145-10
8	59.5	27.4	595	1	US-08-171-718-16
9	59.5	27.4	595	3	US-08-478-087-16
10	59.5	27.4	596	1	US-08-179-738-2
11	59.5	27.4	596	2	US-08-179-738-3
12	59.5	27.4	596	2	US-08-628-145-2
13	59.5	27.4	596	2	US-08-628-145-3
14	54.5	25.1	323	3	US-09-132-619-4
15	54.5	25.1	323	3	US-09-382-803B-4
16	54.5	25.1	323	4	US-09-510-654-4
17	54	24.9	529	2	US-08-933-227-4
18	53.5	24.7	2842	1	US-07-741-940-7
19	53.5	24.7	2842	1	US-08-289-548A-7
20	53.5	24.7	2842	1	US-08-452-654-7
21	53.5	24.7	2843	1	US-07-741-940-2
22	53.5	24.7	2843	1	US-08-289-548A-2
23	53.5	24.7	2843	1	US-08-452-654-2
24	53.5	24.7	2843	1	US-08-452-655B-2
25	53.5	24.7	2843	1	US-08-452-655B-7
26	53.5	24.7	2843	2	US-08-370-235A-2
27	53.5	24.7	2843	3	US-08-450-582-2

28	53.5	24.7	2843	3	US-08-450-582-7	Sequence 7, Appli
29	53.5	24.7	2973	2	US-08-821-355A-7	Sequence 7, Appli
30	53.5	24.7	2973	2	US-09-003-687A-7	Sequence 7, Appli
31	53.5	24.7	2973	4	US-09-136-605-7	Sequence 7, Appli
32	53	24.4	682	4	US-08-982-785A-9	Sequence 9, Appli
33	50.5	23.3	2475	4	US-09-413-814-48	Sequence 48, Appli
34	50	23.0	2016	4	US-09-634-920-4	Sequence 4, Appli
35	49.5	22.8	1004	4	US-08-916-352-2	Sequence 2, Appli
36	49	22.6	160	2	US-08-602-208-4	Sequence 4, Appli
37	49	22.6	4545	2	US-08-804-227C-14	Sequence 14, Appli
38	49	22.6	4550	2	US-08-804-227C-8	Sequence 8, Appli
39	49	22.6	4550	2	US-08-804-198-2	Sequence 3, Appli
40	48.5	22.4	315	3	US-09-099-676-3	Sequence 3, Appli
41	48.5	22.4	315	4	US-09-565-910-3	Sequence 3, Appli
42	48	22.1	295	2	US-08-464-517-20	Sequence 20, Appli
43	48	22.1	295	2	US-08-246-361A-20	Sequence 20, Appli
44	48	22.1	295	3	US-08-463-772-20	Sequence 20, Appli
45	48	22.1	295	5	PCT-US93-05000-20	Sequence 20, Appli

ALIGNMENTS

RESULT 1
US-08-246-583-3
; Sequence 3, Application US/08246583
; Patent No. 5750394
; GENERAL INFORMATION:
; APPLICANT: Palese, Peter
; TITLE OF INVENTION: IDENTIFICATION AND USE OF ANTIVIRAL
; TITLE OF INVENTION: COMPOUNDS THAT INHIBIT INTERACTION OF HOST CELL PROTEINS
; TITLE OF INVENTION: AND VIRAL PROTEINS REQUIRED FOR VIRAL REPLICATION
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/246,583
; FILING DATE: 20-MAY-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 6923-040
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 542 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-246-583-3

Query Match 28.1%; Score 61; DB 1; Length 542;
Best Local Similarity 30.3%; Pred. No. 0.45;
Matches 10; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

Oy 8 IVEAGMGALGLHLTDPQRLVQNCWLTLRLNS 40

COUNTRY: U.S.A.
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0,
CURRENT APPLICATION DATA:

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/628,145
FILING DATE: 04-APR-1996

APPLICATION NUMBER: US 08/628,142
FILING DATE: 04-APR-1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/179,738

CLASSIFICATION: 1330
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/179,738
FILING DATE: 10-JAN-1994
ATTORNEY/AGENT INFORMATION:

APPLICATION NUMBER: 08 00/215798
FILING DATE: 10-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Robins, Roberta L.
REGISTRATION NUMBER: 33,208

ATTORNEY/AGENCY INFORMATION:
NAME: Robins, Roberta L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 5998-001
TELECOMMUNICATION INFORMATION:

REGISTRATION NUMBER: 5998-001
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231

TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ. ID NO: 7:
SEQUENCE CHARACTERISTICS:

INFORMATION FOR SEQ. ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 584 amino acids
 TYPE: amino acid
 TOPOLOGY: linear

```

LENGTH: 584 amino acids
TYPE: amino-acid :
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: N-terminal

```

TOPOLOGY: linear
 MOLECULE TYPE: protein
 FRAGMENT TYPE: N-terminal
 ORIGINAL SOURCE:
 ORGANISM: murine

ORIGINAL SOURCE: murine
-628-145-7

9-628-145-7

Very Match	27.4%	Score 5
St. Local Similarity	45.2%	Pred. N
chès 14; Conservative		4; Mism

13 GMQALGHLTPSQRLVQNCL-WT-LRNLS
4: |||||: || || || : : : :
14; Conservative 4; Mismatch
15; Local Similarity: 45.2%; Phe.
16; Local Similarity: 45.2%; Phe.

13 GMQALGLHLTPSQRLVQNCL--WT-LRNLS
 235 GVDALGLHIYDPENRLTPKISFPWNEIRNIS

235 GVDALGLHIYDPENRLTPKISFPWNEIRNIS

3-179-738-5
Sequence 5, Application US/08179738

3-179-738-5
quence 5, Application US/08179738
cent No. 5578462
GENERAL INFORMATION:
APPLICANT: Seizinger, Bernd R.

GENERAL INFORMATION:
 APPLICANT: Seizinger, Bernd R.
 APPLICANT: Kley, Nikolai A.
 APPLICANT: Bianchi, Albert B.

APPLICANT: Seizinger, Bernd K.
 APPLICANT: Kley, Nikolai A.
 APPLICANT: Bianchi, Albert B.
 TITLE OF INVENTION: No. 5578462el
 NUMBER OF SEQUENCES: '26

TITLE OF INVENTION: NO. 5578462el N
 NUMBER OF SEQUENCES: 26
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Reed & Robins

CORRESPONDENCE ADDRESS:
ADDRESSEE: Reed & Robins
STREET: 635 Bryant Street
CITY: Palo Alto
STATE: California

STREET: 635 Bryant Street
CITY: Palo Alto
STATE: California
COUNTRY: U.S.A.
ZIP: 94301

STATE: California
COUNTRY: U.S.A.
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```

- ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0,
CURRENT APPLICATION DATA:

'SOFTWARE: PatentIn Release #1.0.
CURRENT APPLICATION DATA:
'APPLICATION NUMBER: US/08/179,730
FILING DATE: 10-JAN-1994
CLASSIFICATION: 530

APPLICATION NUMBER: US/08/179,733
FILING DATE: 10-JAN-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Robins, Roberta L.

CLASSIFICATION: #330
ATTORNEY/AGENT INFORMATION:
NAME: Robins, Roberta L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 5998-00

NAME: ROELINS, ROBERT L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 5998-00
TELECOMMUNICATION INFORMATION:

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 591 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE: murine
ORGANISM: murine
US-08-179-738-5

Query Match 27.4%; Score 59.5; DB 1; Length 591;
Best Local Similarity 45.2%; Pred. No. 0.85;
Matches 14; Conservative 4; Mismatches 10; Indels 3; Gaps 2;

Qy 13 GQALGLHLTPSQRVQNCL--WT-LRNLS 40
Db 235 GVDALGLHYDPENRLTPKISFPWNEIRNIS 265

RESULT 5
US-08-179-738-10
Sequence 10, Application US/08179738
Patent No. 5578462
GENERAL INFORMATION:
APPLICANT: Seizinger, Bernd R.
APPLICANT: Kley, Nikolai A.
APPLICANT: Bianchi, Albert B.
TITLE OF INVENTION: NO. 5578462el NF2 Isoforms
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Reed & Robins
STREET: 635 Bryant Street
CITY: Palo Alto
STATE: California
COUNTRY: U.S.A
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/179,738
FILING DATE: 10-JAN-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Robins, Roberta L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 5998-0017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 591 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE: Homo sapien
ORGANISM: Homo sapien
US-08-179-738-10

Query Match 27.4%; Score 59.5; DB 1; Length 591;
Best Local Similarity 45.2%; Pred. No. 0.85;
Matches 14; Conservative 4; Mismatches 10; Indels 3; Gaps 2;

Qy 13 GQALGLHLTPSQRVQNCL--WT-LRNLS 40

Db 235 GVDALGLHYDPENRLTPKISFPWNEIRNIS 265

RESULT 6
US-08-628-145-5
Sequence 5, Application US/08628145
Patent No. 5872214
GENERAL INFORMATION:
APPLICANT: Seizinger, Bernd R.
APPLICANT: Kley, Nikolai A.
APPLICANT: Bianchi, Albert B.
TITLE OF INVENTION: NO. 5872214el NF2 Isoforms
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Reed & Robins
STREET: 635 Bryant Street
CITY: Palo Alto
STATE: California
COUNTRY: U.S.A
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/628,145
FILING DATE: 04-APR-1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/179,738
FILING DATE: 10-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Robins, Roberta L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 5998-0017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 591 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE: murine
ORGANISM: murine
US-08-628-145-5

Query Match 27.4%; Score 59.5; DB 2; Length 591;
Best Local Similarity 45.2%; Pred. No. 0.85;
Matches 14; Conservative 4; Mismatches 10; Indels 3; Gaps 2;

Qy 13 GQALGLHLTPSQRVQNCL--WT-LRNLS 40
Db 235 GVDALGLHYDPENRLTPKISFPWNEIRNIS 265

RESULT 7
US-08-628-145-10
Sequence 10, Application US/08628145
Patent No. 5872214
GENERAL INFORMATION:
APPLICANT: Seizinger, Bernd R.
APPLICANT: Kley, Nikolai A.
APPLICANT: Bianchi, Albert B.
TITLE OF INVENTION: NO. 5872214el NF2 Isoforms
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Reed & Robins

STREET: 635 Bryant Street
CITY: Palo Alto
STATE: California
COUNTRY: U.S.A
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/628,145
FILING DATE: 04-APR-1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/179,738
FILING DATE: 10-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Robins, Roberta L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 5998-0017
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 591 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapien
US-08-628-145-10

Query Match 27.4%; Score 59.5; DB 2; Length 591;
Best Local Similarity 45.2%; Pred. No. 0.85;
Matches 14; Conservative 4; Mismatches 10; Indels 3; Gaps 2;

QY 13 GMAQLGHLTDPSQRLVQNCL--WT-LRNLS 40
I: |||||: || || | :|||:
Db 235 GVDALGLHYDPENRLTPKISFPWNEIRNIS 265

RESULT 8
US-08-171-718-16
Sequence 16, Application US/08171718
Patent No. 5707863
GENERAL INFORMATION:
APPLICANT: Trofatter, James A.
APPLICANT: MacCollin, Mia M.
TITLE OF INVENTION: Tumor Suppressor Gene Merlin and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 120
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, N.W., Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/171,718
FILING DATE: 22-DEC-1993
CLASSIFICATION: 436
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/108,808
FILING DATE: 19-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/022,034
FILING DATE: 25-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/026,063
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Anne
REGISTRATION NUMBER: 36,463
REFERENCE/DOCKET NUMBER: 0609.3850003
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 595 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-171-718-16

Query Match 27.4%; Score 59.5; DB 1; Length 595;
Best Local Similarity 45.2%; Pred. No. 0.85;
Matches 14; Conservative 4; Mismatches 10; Indels 3; Gaps 2;

QY 13 GMAQLGHLTDPSQRLVQNCL--WT-LRNLS 40
I: |||||: || || | :|||:
Db 235 GVDALGLHYDPENRLTPKISFPWNEIRNIS 265

RESULT 9
US-08-478-087-16
Sequence 16, Application US/08478087
Patent No. 6077685
GENERAL INFORMATION:
APPLICANT: Trofatter, James A.
APPLICANT: MacCollin, Mia M.
TITLE OF INVENTION: Tumor Suppressor Gene Merlin and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 120
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, N.W., Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/478,087
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/171,718
FILING DATE: 22-DEC-1993
APPLICATION NUMBER: US 08/108,808
FILING DATE: 19-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/022,034
FILING DATE: 25-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/026,063
FILING DATE: 04-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Anne

REGISTRATION NUMBER: 36,463
REFERENCE/DOCKET NUMBER: 0609.3850003
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 595 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-478-087-16

Query Match 27.4%; Score 59.5; DB 3; Length 595;
Best Local Similarity 45.2%; Pred. No. 0.86; Mismatches 4; Indels 3; Gaps 2;

QY 13 GMDALGLHLTPSQRVLQVNCN--WT-LRNLS 40
Db 235 GVDALGLHYDPENRLTPKISFPWNEIRNIS 265

RESULT 10
US-08-179-738-2
Sequence 2, Application US/08179738
Patent No. 5578462
GENERAL INFORMATION:
APPLICANT: Seizinger, Bernd R.
APPLICANT: Kley, Nikolai A.
APPLICANT: Bianchi, Albert B.
TITLE OF INVENTION: No. 5578462el NF2 Isoforms
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Reed & Robins
STREET: 635 Bryant Street
CITY: Palo Alto
STATE: California
COUNTRY: U.S.A.
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/179,738
FILING DATE: 10-JAN-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Robins, Roberta L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 5998-0017
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 596 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: murine
US-08-179-738-2

Query Match 27.4%; Score 59.5; DB 1; Length 596;
Best Local Similarity 45.2%; Pred. No. 0.86; Mismatches 4; Indels 3; Gaps 2;

QY 13 GMDALGLHLTPSQRVLQVNCN--WT-LRNLS 40

Db 235 GVDALGLHYDPENRLTPKISFPWNEIRNIS 265

RESULT 11
US-08-179-738-3
Sequence 3, Application US/08179738
Patent No. 5578462
GENERAL INFORMATION:
APPLICANT: Seizinger, Bernd R.
APPLICANT: Kley, Nikolai A.
APPLICANT: Bianchi, Albert B.
TITLE OF INVENTION: No. 5578462el NF2 Isoforms
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Reed & Robins
STREET: 635 Bryant Street
CITY: Palo Alto
STATE: California
COUNTRY: U.S.A.
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/179,738
FILING DATE: 10-JAN-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Robins, Roberta L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 5998-0017
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 596 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapien
US-08-179-738-3

Query Match 27.4%; Score 59.5; DB 1; Length 596;
Best Local Similarity 45.2%; Pred. No. 0.86; Mismatches 4; Indels 3; Gaps 2;

QY 13 GMDALGLHLTPSQRVLQVNCN--WT-LRNLS 40
Db 235 GVDALGLHYDPENRLTPKISFPWNEIRNIS 265

RESULT 12
US-08-628-145-2
Sequence 2, Application US/08628145
Patent No. 5872214
GENERAL INFORMATION:
APPLICANT: Seizinger, Bernd R.
APPLICANT: Kley, Nikolai A.
APPLICANT: Bianchi, Albert B.
TITLE OF INVENTION: No. 5872214el NF2 Isoforms
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Reed & Robins
STREET: 635 Bryant Street
CITY: Palo Alto
STATE: California

COUNTRY: U.S.A
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/628.145
FILING DATE: 04-APR-1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/179,738
FILING DATE: 10-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Robins, Roberta L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 5998-0017
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3323
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 596 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: murine
J-08-628-145-2

Query Match	27.4%;	Score 59.5;	DB 2;	Length 596;
Best Local Similarity	45.2%;	Pred. No. 0.86;		
Matches 14: Conservative	4;	Mismatches 10;	Indels	

QY 13 GMAALGLHLTPDPSQRLVQNCL--WT-LRNLS 40
| : | | | | : | | | | : | | : | | : | |
Db 235 GVDALGLHIYDPENRLTPKISFPWNEIRNIS 265

```

RESULT      13
US-08-628-145-3
; Sequence 3, Application US/08G28145
; Patent No. 5872214
; GENERAL INFORMATION:
; APPLICANT: Seizinger, Bernd R.
; APPLICANT: Kleys, Nikolai A.
; APPLICANT: Bianchi, Albert B.
; TITLE OF INVENTION: No. 5872214el NF2 Isoforms
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reed & Robins
; STREET: 635 Bryant Street
; CITY: Palo Alto
; STATE: California
; COUNTRY: U.S.A
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/628,145
; FILING DATE: 04-APR-1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/179,738
; FILING DATE: 10-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Robins, Roberta L.

```

REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 5998-0017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ. ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 596 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapien
US-08-628-145-3

Query Match 27.4%; Score 59.5; DB 2; Length 596;
Best Local Similarity 45.2%; Pred. NO. 0.86;
Matches 14; Conservative 4; Mismatches 10; Indels

QY 13 GMQALGLHLTDPSQRLVQNCL--WT-LRNLS 40
 | : | | | | : | | | : | | : | : |
Db 235 GVDALGLHIYDPENRLTPKISFPWNEIRNIS 265

RESULT 14.
US-09-132-619-4
Sequence 4, Application US/09132619B
Patent No. 5958697
GENERAL INFORMATION:
APPLICANT: Shan, Bel
APPLICANT: Nitta, Masahiro
TITLE OF INVENTION: Cyp7 Promoter-Binding Factors
FILE REFERENCE: T97-013
CURRENT APPLICATION NUMBER: US/09/132.619B
CURRENT FILING DATE: 1998-08-11
EARLIER APPLICATION NUMBER: 60/067,708
EARLIER FILING DATE: 1997-12-08
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 323
TYPE: PRT
ORGANISM: human
US-09-132-619-4

	Query Match	25.1%	Score 54.5;	DB 2;	Length 323;
Best Local Similarity	32.3%				
Pred. No. 2.3;					
Matches + 10;	Conservative	8;	Mismatches	12;	Indels
					1; Gaps 1;
QY	5' KPAIVEAGGMAQALGLHLTDPQRLVQNC	LWT	35		
		:	:	:	:
		:	:	:	:
		:	:	:	:
		:	:	:	:
DP	136. KKAIFRANGLKLEAMSOYDDQMKLQNC	-WS	165		

RESULT 15:
US-09-282-803B-4
; Sequence 4, Application US/09282803B
; Patent No. 6027901
; GENERAL INFORMATION:
; APPLICANT: Shan, Bei
; APPLICANT: Nitta, Masahiro
; TITLE OF INVENTION: CYP7 Promoter-Binding Factors
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010

Search completed: July 29, 2002, 16:05:10
Job time: 480 sec

1. The first part of the document is a letter from the President of the United States to the Congress, dated January 3, 1862. It is a very long letter, and it contains a great deal of information about the state of the country at that time. It is a very important document, and it is one of the most interesting documents in the collection.

2. The second part of the document is a letter from the President of the United States to the Congress, dated January 3, 1862. It is a very long letter, and it contains a great deal of information about the state of the country at that time. It is a very important document, and it is one of the most interesting documents in the collection.

3. The third part of the document is a letter from the President of the United States to the Congress, dated January 3, 1862. It is a very long letter, and it contains a great deal of information about the state of the country at that time. It is a very important document, and it is one of the most interesting documents in the collection.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 16:09:13 ; Search time 70.63 seconds
(without alignments)
55.779 Million cell updates/sec

Title: US-09-641-104a-9
Perfect score: 217
Sequence: 1 CSSNKPATVEAGGMQALGLH.....TDPQRVLQVNCWTLRNLSD 41
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	217	100.0	781	2 A38973	beta-catenin - hum
2	217	100.0	781	2 S35091	beta-catenin - mou
3	209	96.3	781	2 S35099	beta-catenin - Afr
4	185	85.3	621	2 S35092	plakoglobin - mous
5	185	85.3	738	2 S35093	plakoglobin - Afri
6	185	85.3	744	2 A32905	plakoglobin, desmo
7	179	82.5	843	2 T12689	armadillo segmen
8	176	81.1	820	2 S33794	hypothetical prote
9	173	79.7	806	2 JC4835	beta-catenin - Hyd
10	145.5	67.1	817	2 S33793	hypothetical prote
11	77	35.5	811	2 T43175	armadillo protein
12	70.5	32.5	678	2 T23341	beta-catenin - Cae
13	67.5	31.1	2274	2 T30258	adenomatous polyo
14	61	28.1	542	2 S30884	SRP1 protein - yea
15	60	27.6	522	2 A57319	overgrown hematopo
16	59.5	27.4	591	2 I54368	merlin protein - m
17	59.5	27.4	595	2 S33809	neurofibromin 2 -
18	59.5	27.4	596	2 I68664	merlin - mouse
19	57	26.3	434	2 A97433	icl containing pro
20	57	26.3	434	2 AB2651	isocitrate lyase [
21	56	25.8	660	2 T45588	arm repeat contain
22	55	25.3	296	2 AB0527	ferrichrome-bindin
23	54	24.9	529	2 A56516	nuclear localizati
24	54	24.9	529	2 S57873	pendulin - mouse
25	54	24.9	529	2 S57345	m-importin (nuclea
26	54	24.9	564	2 C69052	conserved hypotnet
27	53.5	24.7	2843	1 RBHUAP	adenomatous polyo
28	53.5	24.7	2845	2 I49505	adenomatous polyo
29	53	24.4	242	2 AI2246	hypothetical prote

30	53	24.4	273	2 B31479	env polyprotein pr
31	53	24.4	298	2 E69532	conserved hypotnet
32	53	24.4	764	2 A45321	protein-glutamine
33	53	24.4	911	2 S28498	gene p120 protein
34	52.5	24.2	924	2 T00518	hypothetical prote
35	52.5	24.2	1012	2 I53172	RAE-28 - mouse
36	52	24.0	116	2 I49555	gene C10 protein -
37	52	24.0	627	2 T47968	dynammin-like prote
38	52	24.0	688	2 AI2516	hypothetical prote
39	52	24.0	760	1 S07896	transcription fact
40	51.5	23.7	404	2 T00750	probable protein p
41	51.5	23.7	415	2 A60086	paired box transcr
42	51.5	23.7	527	2 S32430	abrin-b precursor
43	51.5	23.7	528	2 S32431	abrin-d precursor
44	51.5	23.7	562	2 S16022	abrin-c precursor
45	51	23.5	625	2 G90570	hypothetical prote

ALIGNMENTS

RESULT: 1

A38973

beta-catenin - human

C:Species: Homo sapiens (man)

C>Date: 26-Jan-1996 #sequence_revision 08-Feb-1996 #text_change 08-Oct-1999

C:Accession: A38973; S55356; S31988

R:Huelsken, J.; Birchmeier, W.; Behrens, J.

J. Cell Biol. 127, 2061-2069, 1994

A:Title: E-cadherin and APC compete for the interaction with beta-catenin and the cyt

A:Reference number: A38973; MUID:95105247

A:Accession: A38973

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-781 <HUM>

A:Cross-references: GB:Z19054; NID:g38519; PIDN:CAA79497.1; PID:g38520

A:Experimental source: placenta

R:Nollet, F.; Berx, G.; Molemans, F.; van Roy, F.

submitted to the EMBL Data Library, June 1995

A:Description: H. sapiens beta-catenin mRNA.

A:Reference number: S55356

A:Accession: S55356

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-781 <NOI>

A:Cross-references: EMBL:X87838; NID:g1154853; PIDN:CAA61107.1; PID:g860988

C:Comment: The cytoplasmic protein beta-catenin normally binds to a region in the cyt

microfilament network.

C:Comment: Cellular levels of beta-catenin are regulated in part by the adenomatous p

d cancerous cell growth.

C:Genetics:

A:Gene: GDB:CTNNB1; CTNNB...

A:Cross-references: GDB:I41922; OMIM:116806

A:Map position: 3p22-3p21:3

C:Keywords: apoptosis; carcinogenesis; cell adhesion; cytosol

F:151-676/Region: 40-residue repeats

Query: Match 100.0%; Score 217; DB 2; Length 781;

Best Local Similarity 100.0%; Pred. No. 2.6e-21;

Matches: 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSSNKPATVEAGGMQALGLHLDTPQRVLQVNCWTLRNLSD 41

DB 350 CSSNKPATVEAGGMQALGLHLDTPQRVLQVNCWTLRNLSD 390

RESULT: 2

S35091

beta-catenin - mouse

C:Species: Mus musculus (house mouse)

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999

C:Accession: S35091

R;Butz, S.; Stappert, J.; Weissig, H.; Kemler, R.

Science 257, 1142-1144, 1992

A;Title: Plakoglobin and beta-catenin: distinct but closely related.

A;Reference number: S35091; MUID:92376536

A;Accession: S35091

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-781 <BUT>

A;Cross-references: EMBL:M90364; NID:g192141; PIDN:AAA37280.1; PID:g192142

C;Keywords: cytoskeleton

Query Match 100.0%; Score 217; DB 2; Length 781;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSSNKPATVEAGGQALGLHLTDPSQRLVQNCWLTLRLNSD 41

|||||

Db 350 CSSNKPATVEAGGQALGLHLTDPSQRLVQNCWLTLRLNSD 390

RESULT 3

S35099

beta-catenin - African clawed frog

C;Species: Xenopus laevis (African clawed frog)

C;Date: 10-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 21-Jul-2000

C;Accession: S35099

R;McCrear, P.D.; Turck, C.W.; Gumbiner, B.

Science 254, 1359-1361, 1991

A;Title: A homolog of the armadillo protein in Drosophila (plakoglobin) associated with

A;Reference number: S35099; MUID:92073903

A;Accession: S35099

A;Molecule type: mRNA

A;Residues: 1-781 <MCC>

A;Cross-references: GB:M77013; NID:g214020; PIDN:AAA49670.1; PID:g214021

C;Keywords: cytoskeleton

Query Match 96.3%; Score 209; DB 2; Length 781;
Best Local Similarity 97.8%; Pred. No. 3.2e-20;
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CSSNKPATVEAGGQALGLHLTDPSQRLVQNCWLTLRLNSD 41

|||||

Db 350 CSSNKPATVEAGGQALGLHLTDPSQRLVQNCWLTLRLNSD 390

RESULT 4

S35092

plakoglobin - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 17-Mar-1999

C;Accession: S35092

R;Butz, S.; Stappert, J.; Weissig, H.; Kemler, R.

Science 257, 1142-1144, 1992

A;Title: Plakoglobin and beta-catenin: distinct but closely related.

A;Reference number: S35091; MUID:92376536

A;Accession: S35092

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-621 <BUT>

A;Cross-references: EMBL:M90365

C;Keywords: cytoskeleton

Query Match 85.3%; Score 185; DB 2; Length 621;
Best Local Similarity 87.8%; Pred. No. 4.6e-17;
Matches 36; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CSSNKPATVEAGGQALGLHLTDPSQRLVQNCWLTLRLNSD 41

|||||

Db 217 CPSNKPATVEAGGQALGLHLTDPSQRLVQNCWLTLRLNSD 257

RESULT 5

S35093

plakoglobin - African clawed frog

C;Species: Xenopus laevis (African clawed frog)

C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jul-2000

C;Accession: S35093; S24636

R;Pouquet, B.; Zimbelmann, R.; Franke, W.W.

Differentiation 51, 187-194, 1992

A;Title: Identification of plakoglobin in oocytes and early embryos of Xenopus laevis

A;Reference number: S35093; MUID:93093332

A;Accession: S35093

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-738 <FOU>

A;Cross-references: EMBL:M95593; NID:g214656; PIDN:AAA49931.1; PID:g214657

R;Demarais, A.A.

submitted to the EMBL Data Library, June 1992

A;Reference number: S24636

A;Accession: S24636

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 133-184, 'V', 186-225, 'T', 227-292 <DEM>

A;Cross-references: EMBL:X67078; NID:965252; PID:965253

C;Keywords: cytoskeleton

Query Match 85.3%; Score 185; DB 2; Length 738;
Best Local Similarity 87.8%; Pred. No. 5.5e-17;
Matches 36; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CSSNKPATVEAGGQALGLHLTDPSQRLVQNCWLTLRLNSD 41

|||||

Db 337 CPSNKPATVEAGGQALGLHLTDPSQRLVQNCWLTLRLNSD 377

RESULT 6

A32905

plakoglobin, desmosomal - human

C;Species: Homo sapiens (man)

C;Date: 22-Nov-1989 #sequence_revision 22-Nov-1989 #text_change 29-Aug-1997

C;Accession: A32905

R;Franke, W.W.; Goldschmidt, M.D.; Zimbelmann, R.; Mueller, H.M.; Schiller, D.L.; Cow

Proc. Natl. Acad. Sci. U.S.A. 86, 4027-4031, 1989

A;Title: Molecular cloning and amino acid sequence of human plakoglobin, the common 1

A;Reference number: A32905; MUID:89264555

A;Accession: A32905

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-744 <FRA>

A;Cross-references: GB:M23410

C;Genetics:

A;Gene: GDB:JUP

A;Cross-references: GDB:126565; OMIM:173325

A;Map position: 7pter-7qter

C;Keywords: cytoskeleton

Query Match 85.3%; Score 185; DB 2; Length 744;
Best Local Similarity 87.8%; Pred. No. 5.6e-17;
Matches 36; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CSSNKPATVEAGGQALGLHLTDPSQRLVQNCWLTLRLNSD 41

|||||

Db 340 CPSNKPATVEAGGQALGLHLTDPSQRLVQNCWLTLRLNSD 380

RESULT 7

T12689

armadillo, segment polarity protein - fruit fly (Drosophila melanogaster)

N;Alternate names: protein 86E4.6

C;Species: Drosophila melanogaster

C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 21-Jul-2000

C;Accession: T12689; A31861
R;Ferraz, C.; Vidal, S.; Brun, C.; Bucheton, A.; Demaille, J.G.
submitted to the EMBL Data Library, January 1998
A;Description: Sequencing the distal x chromosome of Drosophila melanogaster.
A;Reference number: Z17572
A;Accession: T12689
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-843 <FER>
A;Cross-references: EMBL:AL021106; NID:el371406; PID:el249776; PIDN:CAA15946.1
A;Experimental source: clone cosmid 63B12
R;Riggelman, B.; Wieschaus, E.; Schedl, P.
Genes Dev. 3, 96-113, 1989
A;Title: Molecular analysis of the Armadillo locus: uniformly distributed transcripts and
A;Reference number: A31861; MUID:89211895
A;Accession: A31861
A;Molecule type: DNA
A;Residues: 1-843 <RIG>
A;Cross-references: EMBL:X54468; NID:g7610; PIDN:CAA38350.1; PID:g7611
C;Genetics:
A;Gene: arm
A;Cross-references: FlyBase:FBgn0000117
A;Introns: 11/1; 557/3; 624/3; 715/3; 748/1
A;Note: 86E4.6
C;Keywords: cytoskeleton

Query Match 82.5%; Score 179; DB 2; Length 843;
 Best Local Similarity 82.9%; Pred. No. 4.2e-16;
 Matches 34; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CSSNKP AIVAGG MQALGLHLTDP SQRLVONCLWTLRLNSD 41
 |||||:|||||:|:| | |||||
 Db 358 CSSNKP AIVAGG MQALAMHLGNMSPRLVONCLWTLRLNSD 398

RESULT 8
 S33794
 hypothetical protein - sea urchin (Tripneustes gratilla)
 C:Species: Tripneustes gratilla
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Aug-1997
 C:Accession: S33794
 R:Rosenthal, E.
 Biochim. Biophys. Acta 1173, 337-341, 1993
 A:Title: Identification of homologues to beta-catenin/plakoglobin/armadillo in two inverted
 A:Reference number: S33793; MUID:93305730
 A:Accession: S33794
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-820 <ROS>

```

Query Match          81.1%; Score 176; DB 2; Length 820;
Best Local Similarity 82.9%; Pred. No. 1e-15;
Matches 34; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY      1  CSSNKPAIVEAGGMOALGLHLTDPQSORLVQNCLWTLRNLSL 41
          | : ||||| ||||| ||||| | ||||| ||||| |||||
Db      366 CHNNKPAIVEAGGMSALGLHLGHHSRLVQNCLWTLRNLSL 406

RESULT      9
JC4835
beta-catenin - Hydra magnipapillata
C:Species: Hydra magnipapillata
C:Date: 15-Aug-1996 #sequence_revision 15-Oct-1996 #text_change 21-Jul-2000
C:Accession: JC4835
R:Hobmayer, E.; Hatta, M.; Fischer, R.; Fujisawa, T.; Holstein, T.W.; Sugiyama, T.
Gene 172, 155-159, 1996
A:Title: Identification of a Hydra homologue of the beta-catenin/plakoglobin/armadillo
A:Reference number: JC4835; MUID:96257271
A:Accession: JC4835
A>Status: nucleic acid sequence not shown

```

A: Molecule type: mRNA
A: Residues: 1-806 <HOB>
A: Cross: references: GB:u36781; NID:g1407600; PID:nAAC47137.1; PID:g1407601
C: Comment: This protein plays roles in cadherin-mediated cell adhesion and in signal
C: Genetics:
A: Gene: betaCtn

```

Query Match      79.7%  Score 173;  DB 2:  Length 806;
Best Local Similarity 82.9%  Pred. No. 2.6e-15;
Matches: 34;  Conservative 2;  Mismatches 5;  Indels 0;  Gaps 0;

QY      1  CSSNKPAIVEAGGQALCLHLTDPQSQRIVQNCLWTLRNLSL 41
db      409  CSSNKPAIVEAGGQCALAHYLSHSTRIIVONCIWTLRNLSL 449

```

RESULT 10
S33793
hypothetical protein - spoonworm (Urechis caupo)
C:Species: Urechis caupo
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Aug-1997
C:Accession: S33793

A: Rosenthal, E.
Biochim. Biophys. Acta 1173, 337-341, 1993
A: Title: Identification of homologues to beta-catenin/plakoglobin/armadillo in two
A: Reference number: S33793; MUID: 93305730
A: Accession: S33793
A: Status: preliminary
A: Molecule type: mRNA
A: Residues: 1-817 <ROS>

Query Match	67.1%;	Score 145.5;	DB 2;	Length 817;
Best Local Similarity	73.2%;	Pred. No. 1.5e-11;		
Matches 30:	Conservative	2: Mismatches	8: Indels	1

QY 1 CSSNKPAIVEAGGMQALGLHLTDPQSQRVLVQNCLWTLRNLSD 41
+ : +
D6 373 CASNKPAIVEAGS-AGPSMHLGHSORLVQNCLWTLRNLSD 412

RESULT 11
T43175
armadillo protein homolog BAR-1 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 18-Feb-2000
C:Accession: T43175; T28813
C:Eisenmann, D.M.; Maloof, J.N.; Simske, J.S.; Kenyon, C.; Kim, S.K.
submitted to the EMBL Data Library, May 1998
A:Description: The beta-catenin homolog BAR-1 and LFT-60 Ras coordinately regulate the
A:Reference number: 222328

```
Query Match      81.1%; Score 176; DB 2; Length 820;
Best Local Similarity 82.9%; Pred. No. 1e-15;
Matches 34; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
```

QY	1	CS	NK	PA	IV	EA	GG	MA	LG	HL	HL	TD	PS	QR	LV	QN	CL	WT	LR	NS	LD	41
			:																			
D6	366	CH	NN	KA	IV	EA	GG	MA	LG <td>HL</td> <td>HL</td> <td>GH</td> <td>HS</td> <td>NR</td> <td>LV <td>QN <td>CL <td>WT <td>LR <td>NS <td>LD</td> <td>406</td> </td></td></td></td></td></td>	HL	HL	GH	HS	NR	LV <td>QN <td>CL <td>WT <td>LR <td>NS <td>LD</td> <td>406</td> </td></td></td></td></td>	QN <td>CL <td>WT <td>LR <td>NS <td>LD</td> <td>406</td> </td></td></td></td>	CL <td>WT <td>LR <td>NS <td>LD</td> <td>406</td> </td></td></td>	WT <td>LR <td>NS <td>LD</td> <td>406</td> </td></td>	LR <td>NS <td>LD</td> <td>406</td> </td>	NS <td>LD</td> <td>406</td>	LD	406

RESULT 9

JC4835

beta-catenin - Hydra magnipapillata

C:Species: Hydra magnipapillata

C:Date: 15-Aug-1996 #sequence_revision 15-Oct-1996 #text_change 21-Jul-2000

C:Accession: JC4835

R:Hobmayer, E.; Hatta, M.; Fischer, R.; Fujisawa, T.; Holstein, T.W.; Sugiyama, T.

Gene 172, 155-159, 1996

A:Title: Identification of a Hydra homologue of the beta-catenin/plakoglobin/armad

A:Reference number: JC4835; MUID:96257271

A:Accession: JC4835

A>Status: nucleic acid sequence not shown

Query Match 27.6%; Score 60; DB 2; Length 522;
Best Local Similarity 30.3%; Pred. No. 3.8;
Matches 10; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 8 IVEAGMQALGLHLTDPSQRLVQNCLWTLRNL 40
:: :||: ||| | :|: ||: |::
Db 320 VIAGGLPRLGLLLQHNKSNIVKEAAWTVSNT 352

Search completed: July 29, 2002, 16:09:14
Job time: 724 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 16:09:53 ; Search time 34.24 Seconds
(without alignments)
46.364 Million cell updates/sec

Title: US-09-641-104A-9
Perfect score: 217
Sequence: 1 CSSNKPATVEAGGQALGLH.....TDPQRVLVQNCILWTLRLNLS 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	217	100.0	781	1 CTNB_HUMAN	P35222 homo sapien
2	217	100.0	781	1 CTNB_MOUSE	Q02248 mus musculu
3	210	96.8	781	1 CTNB_RAT	Q9wu82 rattus norv
4	209	96.3	781	1 CTNB_XENLA	P26233 xenopus lae
5	191	88.0	621	1 ARM_MUSDO	Q02453 musca domes
6	185	85.3	621	1 PLAK_MOUSE	Q02257 mus musculu
7	185	85.3	738	1 PLAK_XENLA	P30998 xenopus lae
8	185	85.3	743	1 PLAK_HUMAN	P14923 homo sapien
9	179	82.5	843	1 ARM_DROME	P18824 drosophila
10	176	81.1	820	1 CTNB_TRIGR	P35223 tripneustes
11	145.5	67.1	818	1 CTNB_URECA	P35224 urechis cau
12	61	28.1	542	1 IMAL_YEAST	Q02821 saccharomyc
13	60	27.6	522	1 IMAL_DROME	P52295 drosophila
14	59.5	27.4	595	1 MERL_HUMAN	P35240 homo sapien
15	59.5	27.4	596	1 MERL_MOUSE	P46662 mus musculu
16	54	24.9	529	1 IMAL_HUMAN	P52292 homo sapien
17	54	24.9	529	1 IMAL_MOUSE	P52293 mus musculu
18	54	24.9	764	1 TGLH_TACTR	Q05187 tachyplesu
19	54	24.9	1211	1 PKPA_HUMAN	Q99569 homo sapien
20	53.5	24.7	2842	1 APC_RAT	P70478 rattus norv
21	53.5	24.7	2843	1 APC_HUMAN	P25054 homo sapien
22	53.5	24.7	2845	1 APC_MOUSE	Q61315 mus musculu
23	53	24.4	273	1 ENV_FLVCL	P21444 feline leuk
24	53	24.4	911	1 CTDL_MOUSE	P30999 mus musculu
25	53	24.4	968	1 CTDL_HUMAN	Q60716 homo sapien
26	52	24.0	116	1 SY06_MOUSE	P27784 mus musculu
27	52	24.0	292	1 CGDI_CHICK	P55169 gallus gall
28	52	24.0	760	1 OCT1_XENLA	P16143 xenopus lae
29	51.5	23.7	415	1 PAX2_MOUSE	P32114 mus musculu
30	51.5	23.7	416	1 PAX2_HUMAN	Q02962 homo sapien
31	51.5	23.7	527	1 ABRB_ABRPR	Q06077 abrus preca
32	51.5	23.7	562	1 ABRB_ABRPR	P28590 abrus preca
33	51	23.5	668	1 ENV_FLVCL	P21443 feline leuk

RESULT

ID	CTNB_HUMAN	STANDARD	PRT	781 AA
AC	P35222	01-FEB-1994 (Rel. 28, Created)		
DT	01-FEB-1994 (Rel. 28, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Beta-catenin.			
GN	CTNNB1 OR CTNNB			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	SEQUENCE FROM N.A.			
RP	TISSUE=Placenta;			
RC	MEDLINE=95105247; PubMed=7806582;			
RX	Huelsen J., Birchmeier W., Behrens J.;			
RT	"E-cadherin and APC compete for the interaction with beta-catenin and the cytoskeleton.";			
RL	J. Cell Biol. 127:2061-2069(1994).			
RN	[2]			
RP	REVIEW			
RA	MEDLINE=20145417; PubMed=10679188;			
RX	Kikuchi A.;			
RT	"Regulation of beta-catenin signaling in the Wnt pathway.";			
RL	Biochem. Biophys. Res. Commun. 268:243-248(2000).			
[3]				
RP	X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 133-664.			
RA	MEDLINE=20578806; PubMed=11136974;			
RA	Graham T.A., Weaver C., Mao F., Kimelman D., Xu W.;			
RL	"Crystal structure of a beta-catenin/Tcf complex.";			
CC	Cell 103:885-896(2000).			
CC	-1- FUNCTION: INVOLVED IN THE REGULATION OF CELL ADHESION AND IN SIGNAL TRANSDUCTION THROUGH THE WNT PATHWAY.			
CC	-1- SUBUNIT: THE CYTOSOLIC FORM BOUND TO AXIN ALLOWS THE STIMULATION OF ITS DEGRADATION: THE CYTOSOLIC FORM ALSO INTERACTS WITH APC AND E-CADHERIN. THE NUCLEAR FORM BINDS TCF/LEF-1 AND MAY ALSO BIND PONTIN52 AND DUPLIN. CAN HETERODIMERIZE WITH ALPHA-CATENIN AND ALSO FORM A TRIMERIC COMPLEX WITH CADHERIN. MEDIATES THE ANCHORAGE OF THE CADHERIN COMPLEX TO THE CORTICAL ACTIN CYTOSKELETON.			
CC	-1- SUBCELLULAR LOCATION: CYTOPLASMIC WHEN IT IS UNSTABILIZED (HIGH LEVEL OF PHOSPHORYLATION). TRANSLOCATES TO THE NUCLEUS WHEN IT IS STABILIZED (LOW LEVEL OF PHOSPHORYLATION).			
CC	-1- PTM: PROBABLY PHOSPHORYLATED BY GSK-3B. WNT-DEPENDENT ACTIVATION OF DVL ANTAGONIZES THE ACTION OF GSK-3B. PHOSPHORYLATION OF BETA-CATENIN IS REDUCED AND THE PROTEIN IS NO LONGER DEGRADED, RESULTING IN ITS ACCUMULATION IN CYTOPLASM.			
CC	-1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.			
CC	-1- SIMILARITY: CONTAINS 12 ARM REPEATS.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial			

entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; X87838; CAA61107.1; -
CC EMBL; Z19054; CAA79497.1; -
DR PIR; S31988; S31988
DR PDB; 1G3J; 17-JAN-01.
DR TRANSFAC; T02872; -
DR MIM; 116806; -
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 12.
DR SMART; SM00185; ARM; 11.
DR PROSITE; PS50176; ARM_REPEAT; 9.
KW Cell adhesion; Cytoskeleton; Structural protein; Nuclear protein;
FT REPEAT 151 191 ARM 1.
FT REPEAT 193 234 ARM 2.
FT REPEAT 235 276 ARM 3.
FT REPEAT 277 318 ARM 4.
FT REPEAT 319 360 ARM 5.
FT REPEAT 361 389 ARM 6.
FT REPEAT 400 441 ARM 7.
FT REPEAT 442 484 ARM 8.
FT REPEAT 489 530 ARM 9.
FT REPEAT 531 571 ARM 10.
FT REPEAT 594 636 ARM 11.
FT REPEAT 637 666 ARM 12.
SQ SEQUENCE 781 AA; 85496 MW; CB78F165A3BEF86E CRC64;

Query Match 100.0%; Score 217; DB 1; Length 781;
Best Local Similarity 100.0%; Pred. No. 1.9e-21;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSSNKPAAIVEAGGQALGLHLTPDSQRLVONCLWTLRLNSD 41
|||||
Db 350 CSSNKPAAIVEAGGQALGLHLTPDSQRLVONCLWTLRLNSD 390

RESULT 2

ID CTNN_MOUSE STANDARD; PRT; 781 AA.
AC Q02248;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Beta-catenin.
GN CTNNB1 OR CATNB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92376536; PubMed=1509266;
RA Butz S., Stappert J., Welssig H., Kemler R.;
RT "Plakoglobin and beta-catenin: distinct but closely related.";
RL Science 257:1142-1144(1992).
RN [2]

RP REVIEW.
RX MEDLINE=20145417; PubMed=10679188;
RA Kikuchi A.;
RT "Regulation of beta-catenin signaling in the Wnt pathway.";
RL Biochem. Biophys. Res. Commun. 268:243-248(2000).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 150-665.
RX MEDLINE=97442350; PubMed=9298899;
RA Huber A.H., Nelson W.J., Weis W.I.;
RT "Three-dimensional structure of the armadillo repeat region of beta-catenin.";
RL Cell 90:871-882(1997).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 118-149 IN COMPLEX WITH

RP CTNNAL.
RX MEDLINE=20337986; PubMed=10882138;
RA Pokutta S., Weis W.I.;
RT "Structure of the dimerization and beta-catenin-binding region of alpha-catenin.";
RL Mol. Cell 5:533-543(2000).
CC -1- FUNCTION: INVOLVED IN THE REGULATION OF CELL ADHESION AND IN SIGNAL TRANSDUCTION THROUGH THE WNT PATHWAY.
CC -1- SUBUNIT: THE CYTOSOLIC FORM BINDS TCF/LEF-1 AND MAY ALSO BIND OF ITS DEGRADATION. THE CYTOSOLIC FORM ALSO INTERACTS WITH APC AND E-CADHERIN. THE NUCLEAR FORM BINDS TCF/LEF-1 AND MAY ALSO BIND PONTIN52 AND DUPLIN. CAN HETERODIMERIZE WITH ALPHA-CATENIN AND ALSO FORM A TRIMERIC COMPLEX WITH CADHERIN. MEDIATES THE ANCHORAGE OF THE CADHERIN COMPLEX TO THE CORTICAL ACTIN CYTOSKELETON.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC WHEN IT IS UNSTABILIZED (HIGH LEVEL OF PHOSPHORYLATION). TRANSLOCATES TO THE NUCLEUS WHEN IT IS STABILIZED (LOW LEVEL OF PHOSPHORYLATION).
CC -1- PTM: PROBABLY PHOSPHORYLATED BY GSK-3B. WNT-DEPENDENT ACTIVATION OF DVL ANTAGONIZES THE ACTION OF GSK-3B. PHOSPHORYLATION OF BETA-CATENIN IS REDUCED AND THE PROTEIN IS NO LONGER DEGRADED, RESULTING IN ITS ACCUMULATION IN CYTOPLASM.
CC -1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -1- SIMILARITY: CONTAINS 12 ARM REPEATS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use, by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M90364; AAA37280.1; -
DR PIR; S35091; S35091.
DR PDB; 2BCT; 15-OCT-97.
DR PDB; 3BCT; 19-NOV-97.
DR PDB; 1DOW; 12-JUL-00.
DR TRANSFAC; T02984; -
DR MGD; MGI:88276; Catnb.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 12.
DR SMART; SM00185; ARM; 11.
DR PROSITE; PS50176; ARM_REPEAT; 9.
KW Cell adhesion; Cytoskeleton; Structural protein; Nuclear protein;
FT REPEAT 151 191 ARM 1.
FT REPEAT 193 234 ARM 2.
FT REPEAT 235 276 ARM 3.
FT REPEAT 277 318 ARM 4.
FT REPEAT 319 360 ARM 5.
FT REPEAT 361 389 ARM 6.
FT REPEAT 400 441 ARM 7.
FT REPEAT 442 484 ARM 8.
FT REPEAT 489 530 ARM 9.
FT REPEAT 531 571 ARM 10.
FT REPEAT 594 636 ARM 11.
FT REPEAT 637 666 ARM 12.
SQ SEQUENCE 781 AA; 85470 MW; D708F170A3FBED6E CRC64;

Query Match 100.0%; Score 217; DB 1; Length 781;
Best Local Similarity 100.0%; Pred. No. 1.9e-21;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CSSNKPAAIVEAGGQALGLHLTPDSQRLVONCLWTLRLNSD 41
|||||
Db 350 CSSNKPAAIVEAGGQALGLHLTPDSQRLVONCLWTLRLNSD 390

RESULT 3
CTNN_MOUSE STANDARD; PRT; 781 AA.
ID CTNN_MOUSE
AC Q9WU82;

DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Beta-catenin.
 GN CTNNB1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN:SPRAGUE-DAWLEY; TISSUE=Testis;
 RX MEDLINE=99428593; PubMed=10497305;
 RA Chung S.S.W., Lee W.M., Cheng C.Y.;
 RT "Study on the formation of specialized inter-Sertoli cell junctions in
 vitro.";
 RL J. Cell. Physiol. 181:258-272(1999).
 CC -!- FUNCTION: INVOLVED IN THE REGULATION OF CELL ADHESION AND IN
 CC SIGNAL TRANSDUCTION THROUGH THE WNT PATHWAY (BY SIMILARITY).
 CC -!- SUBUNIT: THE CYTOSOLIC FORM BOUND TO AXIN ALLOWS THE STIMULATION
 CC OF ITS DEGRADATION. THE CYTOSOLIC FORM ALSO INTERACTS WITH APC AND
 CC E-CADHERIN. THE NUCLEAR FORM BINDS TCF/LEF-1 AND MAY ALSO BIND
 CC PONTIN2 AND DUPLIN. CAN HETERODIMERIZE WITH ALPHA-CATENIN AND
 CC ALSO FORM A TRIMERIC COMPLEX WITH CADHERIN. MEDIATES THE ANCHORAGE
 CC OF THE CADHERIN COMPLEX TO THE CORTICAL ACTIN CYTOSKELETON (BY
 CC SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC WHEN IT IS UNSTABILIZED (HIGH
 CC LEVEL OF PHOSPHORYLATION). TRANSLOCATES TO THE NUCLEUS WHEN IT IS
 CC STABILIZED (LOW LEVEL OF PHOSPHORYLATION) (BY SIMILARITY).
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE TESTIS.
 CC -!- DEVELOPMENTAL STAGE: HIGHLY EXPRESSED AT E30-60 DAY DPC IN THE
 CC TESTIS. REDUCED EXPRESSION AT E90 DAY DPC.
 CC -!- PTM: PROBABLY PHOSPHORYLATED BY GSK-3B. WNT-DEPENDENT ACTIVATION
 CC OF DVL ANTAGONIZES THE ACTION OF GSK-3B. PHOSPHORYLATION OF BETA-
 CC CATENIN IS REDUCED AND THE PROTEIN IS NO LONGER DEGRADED,
 CC RESULTING IN ITS ACCUMULATION IN CYTOPLASM (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
 CC -!- SIMILARITY: CONTAINS 12 ARM REPEATS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF121265; RAD28504.1; -
 CC DR HSP: Q02248; 2BCT.
 CC DR InterPro: IPR000225; Armadillo.
 CC DR Pfam: PF00514; Armadillo_seg; 12.
 CC DR SMART: SM00185; ARM; 11.
 CC DR PROSITE: PS50176; ARM_REPEAT; 9.
 CC Cell adhesion; Cytoskeleton; Structural protein; Nuclear protein;
 KW Repeat.
 KW REPEAT 151 191 ARM 1.
 FT REPEAT 193 234 ARM 2.
 FT REPEAT 235 276 ARM 3.
 FT REPEAT 277 318 ARM 4.
 FT REPEAT 319 360 ARM 5.
 FT REPEAT 361 389 ARM 6.
 FT REPEAT 400 441 ARM 7.
 FT REPEAT 442 484 ARM 8.
 FT REPEAT 489 530 ARM 9.
 FT REPEAT 531 571 ARM 10.
 FT REPEAT 594 636 ARM 11.
 FT REPEAT 637 666 ARM 12.
 SQ SEQUENCE 781 AA; 85454 MW; 9C29186B6DD54B87 CRC64;
 Query Match 96.8%; Score 210; DB 1; Length 781;
 Best Local Similarity 97.6%; Pred. No. 1.7e-20;
 Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1:CSSNKPAAVEAGGMOALGLHLTDPFSORLVONCLWTLRLNSD 41
 DB 350:CSSNKPAAVEAGGMOALGLHLTDPFSORLVONCLWTLRLNSD 390
 RESULT 4
 CTNB_XENLA STANDARD; PRT; 781 AA.
 ID CTNB_XENLA
 AC P26237
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE beta-catenin.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodidae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Brain;
 RX MEDLINE=92073903; PubMed=1962194;
 RA McCrea P.D., Turk C.W., Gumbiner B.M.;
 RT "A homolog of the armadillo protein in Drosophila (plakoglobin)
 associated with E-cadherin.";
 RL Science 254:1359-1361(1991).
 CC -!- FUNCTION: BINDS TO THE CYTOPLASMIC DOMAIN OF THE CELL-CELL
 CC ADHESION MOLECULE E-CADHERIN, AND PERHAPS TO OTHER (MEMBRANE)
 CC PROTEINS. THE ASSOCIATION OF CATENINS TO CADHERINS PRODUCES A
 CC COMPLEX WHICH IS LINKED TO THE ACTIN FILAMENT NETWORK, AND WHICH
 CC SEEMS TO BE OF PRIMARY IMPORTANCE FOR CADHERINS CELL-ADHESION
 CC PROPERTIES. ESSENTIAL COMPONENT OF THE VERTEBRATE WNT-SIGNALING
 CC PATHWAY AND IS REQUIRED FOR FORMATION OF THE EMBRYONIC AXIS.
 CC -!- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
 CC -!- SIMILARITY: CONTAINS 10 ARM REPEATS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M77013; AAA49670.1; -
 CC DR PIR: S35099; S35099.
 CC DR HSP: Q02248; 2BCT.
 CC DR TRANSFAC: T03026; -
 CC DR InterPro: IPR000225; Armadillo.
 CC DR Pfam: PF00514; Armadillo_seg; 12.
 CC DR SMART: SM00185; ARM; 11.
 CC DR PROSITE: PS50176; ARM_REPEAT; 9.
 CC Cell adhesion; Cytoskeleton; Structural protein; Repeat.
 KW REPEAT 141 180 ARM 1.
 FT REPEAT 225 264 ARM 2.
 FT REPEAT 267 306 ARM 3.
 FT REPEAT 351 390 ARM 4.
 FT REPEAT 391 429 ARM 5.
 FT REPEAT 432 473 ARM 6.
 FT REPEAT 479 519 ARM 7.
 FT REPEAT 521 562 ARM 8.
 FT REPEAT 584 623 ARM 9.
 FT REPEAT 625 664 ARM 10.
 SQ SEQUENCE 781 AA; 85449 MW; 3ECD27232239F799 CRC64;
 Query Match 96.3%; Score 209; DB 1; Length 781;
 Best Local Similarity 97.6%; Pred. No. 2.3e-20;
 Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1:CSSNKPAAVEAGGMOALGLHLTDPFSORLVONCLWTLRLNSD 41
 DB 350:CSSNKPAAVEAGGMOALGLHLTDPFSORLVONCLWTLRLNSD 390


```

RESULT 7
PLAK_XENLA STANDARD; PRT; 738 AA.
ID PLAK_XENLA AC P30998;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Junction plakoglobin (Desmoplakin III).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93093332; PubMed=1459359;
RA Fouquet B., Zimbelmann R., Franke W.W.;
RT "Identification of plakoglobin in oocytes and early embryos of
RT Xenopus laevis: maternal expression of a gene encoding a junctional
RT plaque protein."
RL Differentiation 51:187-194(1992).
RN [2]
RP SEQUENCE OF 133-292 FROM N.A.
RX MEDLINE=93012479; PubMed=1397690;
RA de Marais A.A., Moon R.T.;
RT "The armadillo homologs beta-catenin and plakoglobin are
RT differentially expressed during early development of Xenopus
RT laevis."
RL Dev. Biol. 153:337-346(1992).
CC -!- FUNCTION: COMMON JUNCTIONAL PLAQUE PROTEIN. THE MEMBRANE-
CC ASSOCIATED PLAQUES ARE ARCHITECTURAL ELEMENTS IN AN IMPORTANT
CC STRATEGIC POSITION TO INFLUENCE THE ARRANGEMENT AND FUNCTION
CC OF BOTH THE CYTOSKELETON AND THE CELLS WITHIN THE TISSUE. THE
CC PRESENCE OF PLAKOGLOBIN IN BOTH THE DESMOSOMES AND IN THE
CC INTERMEDIATE JUNCTIONS SUGGESTS THAT IT PLAYS A CENTRAL ROLE IN
CC THE STRUCTURE AND FUNCTION OF SUBMEMBRANOUS PLAQUES.
CC -!- SUBUNIT: HOMODIMER.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC IN A SOLUBLE AND MEMBRANE-
CC ASSOCIATED FORM.
CC -!- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -!- SIMILARITY: CONTAINS 10 ARM REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M95593; AAA49931.1; -
DR EMBL; X67078; CAA47463.1; -
DR PIR; S24636; S24636.
DR PIR; S35093; S35093.
DR HSSP; Q02248; 2BCT.
DR InterPro: IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 11.
DR SMART; SM00185; ARM; 10.
DR Cell adhesion; Cytoskeleton; Structural protein; Repeat.
KW REPEAT 128 167 ARM 1.
FT REPEAT 128 167 ARM 1.
FT REPEAT 212 251 ARM 2.
FT REPEAT 254 293 ARM 3.
FT REPEAT 338 377 ARM 4.
FT REPEAT 379 416 ARM 5.
FT REPEAT 419 460 ARM 6.
FT REPEAT 466 506 ARM 7.
FT REPEAT 508 547 ARM 8.
FT REPEAT 570 609 ARM 9.
FT REPEAT 611 650 ARM 10.
FT REPEAT 185 185 I -> V (IN REF. 2).
FT CONFLICT 185
FT CONFLICT 185

FT CONFLICT 226 226 A -> T (IN REF. 2).
SQ SEQUENCE 738 AA; 81711 MW; 569DBE69D08B58 CRC64;

Query: Match 1185.3%; Score 185; DB 1; Length 738;
Best Local Similarity 87.8%; Pred. No. 3.6e-17;
Matches 36; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 11: CSSNKPATVEAGGMOALGHLTDPDSRVQNCVONCLWTLRLNLSLSD 41
DB 337 CPSNKPATVEAGGMOALGKHLTSNPRVLQNCVONCLWTLRLNLSLSD 377

RESULT 18
PLAK_HUMAN STANDARD; PRT; 743 AA.
ID PLAK_HUMAN AC P14923;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Junction plakoglobin (Desmoplakin III).
CN JUP OR DP3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89264555; PubMed=2726765;
RA Franke W.W., Goldschmidt M.D., Zimbelmann R., Mueller H.M.,
RA Schiller D.L., Cowin P.;
RT "Molecular cloning and amino acid sequence of human plakoglobin, the
RT common junctional plaque protein."
RL Proc. Natl. Acad. Sci. U.S.A. 86:4027-4031(1989).
CC -!- FUNCTION: COMMON JUNCTIONAL PLAQUE PROTEIN. THE MEMBRANE-
CC ASSOCIATED PLAQUES ARE ARCHITECTURAL ELEMENTS IN AN IMPORTANT
CC STRATEGIC POSITION TO INFLUENCE THE ARRANGEMENT AND FUNCTION
CC OF BOTH THE CYTOSKELETON AND THE CELLS WITHIN THE TISSUE. THE
CC PRESENCE OF PLAKOGLOBIN IN BOTH THE DESMOSOMES AND IN THE
CC INTERMEDIATE JUNCTIONS SUGGESTS THAT IT PLAYS A CENTRAL ROLE IN
CC THE STRUCTURE AND FUNCTION OF SUBMEMBRANOUS PLAQUES.
CC -!- SUBUNIT: HOMODIMER.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC IN A SOLUBLE AND MEMBRANE-
CC ASSOCIATED FORM.
CC -!- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -!- SIMILARITY: CONTAINS 8 ARM REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M23410; AAA64895.1; -
DR PIR; A32905; A32905
DR HSSP; Q02248; 2BCT.
DR MIM; 173325; -
DR InterPro: IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 10.
DR SMART; SM00185; ARM; 8.
DR PROSITE; PS00176; ARM_REPEAT; 8.
DR Cell adhesion; Cytoskeleton; Structural protein; Repeat.
KW INIT_MET 0
FT REPEAT 141 179 ARM 1.
FT REPEAT 183 224 ARM 2.
FT REPEAT 225 252 ARM 3.
FT REPEAT 308 351 ARM 4.
FT REPEAT 389 430 ARM 5.
FT REPEAT 431 473 ARM 6.
FT REPEAT 478 521 ARM 7.
FT REPEAT 582 624 ARM 8.

```


CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----

DR EMBL; X54468; CAA38350.1; .
 DR EMBL; AF001213; AAB58731.1; .
 DR EMBL; AE003422; AAF45688.1; ALT_INIT.
 DR EMBL; AL021106; CAA15946.1; .
 DR EMBL; AL021086; CAA15946.1; JOINED.
 DR EMBL; AL021086; CAA15935.1; .
 DR EMBL; AL021106; CAA15935.1; JOINED.
 DR PTR; A31861; A31861.
 DR HSP; Q02248; IDOW.
 DR TRANSFAC; T02977; .
 DR FlyBase; FBgn000117; arm.
 DR InterPro; IPR000225; Armadillo.
 DR Pfam; PF00514; Armadillo_seg; 12.
 DR SMART; SM00185; ARM; 11.
 DR PROSITE; PS50176; ARM_REPEAT; 9.
 KW Developmental protein; Segmentation polarity protein; Repeat;
 KW Cell adhesion; Cytoskeleton; Structural protein; Phosphorylation;
 KW Alternative splicing.
 FT DOMAIN 1 158 ASP/GLU-RICH (ACIDIC).
 FT REPEAT 159 200 ARM 1.
 FT REPEAT 201 242 ARM 2.
 FT REPEAT 243 284 ARM 3.
 FT REPEAT 285 326 ARM 4.
 FT REPEAT 327 368 ARM 5.
 FT REPEAT 369 410 ARM 6.
 FT REPEAT 411 449 ARM 7.
 FT REPEAT 450 496 ARM 8.
 FT REPEAT 497 538 ARM 9.
 FT REPEAT 539 584 ARM 10.
 FT REPEAT 585 608 ARM 11.
 FT REPEAT 609 647 ARM 12.
 FT REPEAT 648 689 ARM 13 (INCOMPLETE).
 FT DOMAIN 690 843 ASP/GLU-RICH (ACIDIC).
 FT VARSPLIC 718 LGPEEAYEGLYGQPPSVSHSGHGFHQGYDLPIDSMQ
 FT FT GLTSSVPVGGGGAGGAPGNGGAVGGAGGGGNGNIGAIPPSGA
 FT FT PTSPYMDMDVGEIDAGALNFDLDAMPTPPDNNNNLAAYD
 FT FT TDC -> ILXQ (IN NEURAL ISOFORM).
 SQ SEQUENCE 843 AA; 91152 MW; 40DAD6FB83163049 CRC64;

Query Match 82.5%; Score 179; DB 1; Length 843;
 Best Local Similarity 82.9%; Pred. No. 2.7e-16;
 Matches 34; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CSSNKPAAVEAGGQALGLHLTDPQSRLVQNCWLTLRLNSD 41
 DB 358 CSSNKPAAVDAGGQALAMHLGNNSPRLVQNCWLTLRLNSD 398
 RESULT 10
 CTNB_TRIGR
 ID CTNB_TRIGR STANDARD; PRT; 820 AA.
 AC P35223;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Beta-catenin.
 OS Tripneustes gratilla (Hawaiian sea urchin).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoidae; Euechinoidea; Echinacea; Temnopleuroidea; Toxopneustidae;
 OC Tripneustes
 OX NCBI_TaxID=7673;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93305730; PubMed=8318544;
 RA Rosenthal E.T.;
 RT "Identification of homologues to beta-catenin/plakoglobin/armadillo
 in two invertebrates, Urechis caupo and Tripneustes gratilla";
 RL Biochim. Biophys. Acta 1173:337-341(1993).
 CC -1- SIMILARITY: BELONGS TO THE CYTOPLASMIC DOMAIN OF THE CELL-CELL
 CC -1- FUNCTION: BINDS TO THE ACTIN FILAMENT NETWORK, AND WHICH
 CC COMPLEX WHICH IS LINKED TO THE ACTIN FILAMENT NETWORK, AND WHICH
 CC SEEMS TO BE OF PRIMARY IMPORTANCE FOR CADHERINS CELL-ADHESION
 CC PROPERTIES.
 CC -1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 7 ARM REPEATS.

CC ADHESION MOLECULE E-CADHERIN, AND PERHAPS TO OTHER (MEMBRANE)
 CC PROTEINS. THE ASSOCIATION OF CATENINS TO CADHERINS PRODUCES A
 CC COMPLEX WHICH IS LINKED TO THE ACTIN FILAMENT NETWORK, AND WHICH
 CC SEEMS TO BE OF PRIMARY IMPORTANCE FOR CADHERINS CELL-ADHESION
 CC PROPERTIES.
 CC -1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 10 ARM REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use; by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----
 CC EMBL; L10354; AAA30089.1; .
 DR PIR; S33794; S33794.
 DR HSP; Q02248; IDOW.
 DR InterPro; IPR000225; Armadillo.
 DR Pfam; PF00514; Armadillo_seg; 12.
 DR SMART; SM00185; ARM; 12.
 DR PROSITE; PS50176; ARM_REPEAT; 9.
 KW Cell adhesion; Cytoskeleton; Structural protein; Repeat.
 FT REPEAT 157 196 ARM 1.
 FT REPEAT 199 239 ARM 2.
 FT REPEAT 241 280 ARM 3.
 FT REPEAT 283 322 ARM 4.
 FT REPEAT 367 405 ARM 5.
 FT REPEAT 406 445 ARM 6.
 FT REPEAT 448 489 ARM 7.
 FT REPEAT 495 535 ARM 8.
 FT REPEAT 603 642 ARM 9.
 FT REPEAT 644 683 ARM 10.
 SQ SEQUENCE 820 AA; 89361 MW; 57255E0F57795FD3 CRC64;
 Query Match 81.1%; Score 176; DB 1; Length 820;
 Best Local Similarity 82.9%; Pred. No. 6.6e-16;
 Matches 34; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 QY 1 CSSNKPAAVEAGGQALGLHLTDPQSRLVQNCWLTLRLNSD 41
 DB 366 CHNNKPAAVEAGGMSALGLHLGHHSNRLVQNCWLTLRLNSD 406
 RESULT 11
 CTNB_URECA
 ID CTNB_URECA STANDARD; PRT; 818 AA.
 AC P35224;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Beta-catenin.
 OS Urechis caupo (Innkeeper worm) (Spoonworm).
 OC Eukaryota; Metazoa; Echiura; Xenopneusta; Urechidae; Urechis.
 OX NCBI_TaxID=6431;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93305730; PubMed=8318544;
 RA Rosenthal E.T.;
 RT "Identification of homologues to beta-catenin/plakoglobin/armadillo
 in two invertebrates, Urechis caupo and Tripneustes gratilla";
 RL Biochim. Biophys. Acta 1173:337-341(1993).
 CC -1- FUNCTION: BINDS TO THE CYTOPLASMIC DOMAIN OF THE CELL-CELL
 CC ADHESION MOLECULE E-CADHERIN, AND PERHAPS TO OTHER (MEMBRANE)
 CC PROTEINS. THE ASSOCIATION OF CATENINS TO CADHERINS PRODUCES A
 CC COMPLEX WHICH IS LINKED TO THE ACTIN FILAMENT NETWORK, AND WHICH
 CC SEEMS TO BE OF PRIMARY IMPORTANCE FOR CADHERINS CELL-ADHESION
 CC PROPERTIES.
 CC -1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 7 ARM REPEATS.

DR PROSITE; PS00661; BAND_41_2; 1.
DR PROSITE; PS00057; BAND_41_3; 1.
KW Structural protein; Cytoskeleton; Anti-oncogene; Alternative splicing.
FT DOMAIN 74 241
FT VARSPLIC 581 596
FT
FT CONFLICT 475 475
FT CONFLICT 554 554
FT CONFLICT 570 570
SQ SEQUENCE 596 AA; 69691 MW; 8D06E3B07D2C5851 CRC64;
BAND 4.1-LIKE.
LTLOSASRVAFFEEEL -> PQAQGRRPICI (IN
SECOND ISOFORM). 2).

Query Match 27.4%; Score 59.5; DB 1; Length 596;
Best Local Similarity 45.2%; Pred. No. 2.1;
Matches 14; Conservative 4; Mismatches 10; Indels 3; Gaps 2;

QY 13 GMDALGLHLDPSQRLVQNCI--WT-LRNLS 40
|: |||||: || || | :||:|
Db 235 GVDALGLHIYDPENRLTPKISFFWNEIRNIS 265

Search completed: July 29, 2002, 16:09:54
Job time: 759 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 16:07:55 ; Search time 158.47 Seconds
(without alignments)
27.336 Million cell updates/sec

Title: US-09-641-104A-10

Perfect score: 194

Sequence: 1 AATKQEGMEGLLTQLVLLGSDDDINVTCAAGILSNLTC 39

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq_032802.*

```

1: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	194	100.0	39	20	AA1980
2	194	100.0	39	20	AA1981
3	194	100.0	781	21	AA1982
4	194	100.0	781	21	AA1983
5	194	100.0	781	22	AA1984
6	194	100.0	781	22	AA1985
7	194	100.0	781	22	AA1986
8	194	100.0	781	22	AA1987
9	155	79.9	840	22	AA1988
10	155	79.9	840	22	AA1989
11	155	79.9	840	22	AA1990

12	128	66.0	700	12	AA11354	Cellular Receptor
13	57	29.4	228	22	ABB64252	Drosophila melanog
14	55	28.4	2931	22	ABB68229	Drosophila melanog
15	55	28.4	3023	17	AA194462	Hepatitis C virus
16	54	27.8	3440	21	AA133301	Arabidopsis thalia
17	54	27.8	442	21	AA133300	Arabidopsis thalia
18	54	27.8	458	21	AA133299	Arabidopsis thalia
19	53	27.3	497	22	AA132203	Novel human secret
20	52	26.8	452	11	AA135689	LXIII sequence for
21	52	26.8	452	21	AA148459	Arabidopsis thalia
22	52	26.8	555	21	AA148458	Arabidopsis thalia
23	52	26.8	606	22	ABB69121	Drosophila melanog
24	51.5	26.5	343	22	AA186268	C. glutamicum pfka
25	51.5	26.5	346	22	AA179269	Corynebacterium gl
26	51.5	26.5	357	22	AA191129	C glutamicum prote
27	51.5	26.5	401	21	AA181503	Streptococcus pneu
28	51.5	26.5	465	22	ABB64710	Drosophila melanog
29	51.5	26.5	578	12	AA100098	Alpha subunit of A
30	51	26.3	178	18	AA183894	Flea saliva protei
31	51	26.3	178	19	AA182368	Flea saliva protei
32	51	26.3	345	21	AA168665	Amino acid sequenc
33	51	26.3	345	22	AA131750	Amino acid sequenc
34	51	26.3	345	22	AA167193	Mastadenovirus 45
35	51	26.3	502	20	AA173956	Protoporphyrinogen
36	50.5	26.0	119	22	AA131622	Amino acid sequenc
37	50.5	26.0	119	22	AA131624	Amino acid sequenc
38	50.5	26.0	192	22	ABB69449	Drosophila melanog
39	50	25.8	303	22	AA133897	Staphylococcus aur
40	50	25.8	306	22	AA136707	Staphylococcus aur
41	50	25.8	346	21	AA168647	Amino acid sequenc
42	50	25.8	346	22	AA131732	Amino acid sequenc
43	50	25.8	346	22	AA167175	Mastadenovirus 23
44	50	25.8	453	18	AA161328	Human host cell pr
45	50	25.8	453	22	AA170386	Human host cell pr

ALIGNMENTS

```

RESULT 1
AA133221
ID AAY33221 standard; peptide; 39 AA.
XX
AC AAY33221;
XX
DT 18-NOV-1999 (first entry)
XX
DE Human beta-catenin protein armadillo repeat arm7 fragment.
XX
KW Beta-catenin; human; armadillo repeat; treatment; human disease; LEF-1;
KW modulator; interaction domain; transcription factor; TCF-4; oncogenic;
KW tumor suppressor gene; conductin; E-cadherin; screening; carcinoma; APC;
KW Wnt signalling pathway; tumor development; anti-oncogenic; melanoma;
KW organ regeneration; tissue regeneration; hair growth.
XX
OS Homo sapiens.
XX
PN DE19909251-Al.
XX
PD 26-AUG-1999.
XX
PF 22-FEB-1999; 99DE-1009251.
XX
PR 21-FEB-1998; 98DE-1007390.
XX
PA (DELE-) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.
XX
PI Birchmeier W, Von Kries J;
XX
DR WPI; 1999-470389/40.
XX
PT Agents for treating human diseases, particularly cancer, modulate
PT interaction of beta-catenin with transcription factors or tumor

```

PT suppressor gene products
 XX Disclosure; Page 7; 16pp; German.
 PS This invention describes a novel agent (A) for treating human disease
 XX which is based on substances (I) that modulate (inhibit or promote) the
 CC interaction of beta-catenin with transcription factors or products of
 CC tumor suppressor genes. The invention also describes (a) peptides (II)
 CC comprising part of the LEF-1/TCF-4 transcription factors, or their
 CC variants and mutants; (b) peptides and related molecules (III) from the
 CC armadillo domain (arm units 3-8) of beta-catenin, and mutants of the
 CC complete beta-catenin molecule, that include at least one of the
 CC specific interaction domains for LEF-1, TCF-4, APC, conductin or
 CC E-cadherin; and (c) enzyme-linked immunosorbent assay (ELISA) for
 CC screening substance libraries for compounds that modulate interaction of
 CC beta-catenin with LEF-1/TCF, APC, conductin or E-cadherin. Beta-catenin
 CC is a key compound in the Wnt signalling pathway and is involved in
 CC development of tumors. Generally its interaction with LEF-1 or TCF-4 is
 CC oncogenic but interaction with APC, conductin or E-cadherin is
 CC anti-oncogenic. (A) which inhibit interaction are particularly used to
 CC treat tumors, especially carcinoma of the colon and melanoma, but also,
 CC where they promote interaction, to stimulate regeneration of organs and
 CC tissues, specifically hair growth. AAY33217-Y33222 represent human
 CC beta-catenin armadillo repeat fragments described in the method of the
 CC invention.
 XX
 SQ Sequence 39 AA;

Query Match 100.0%; Score 194; DB 20; Length 39;
 Best Local Similarity 100.0%; Pred. No. 4.7e-21;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATKQEGMGLLTGLVQLGSDDDINVTCAAGILSNLTC 39
 |||||
 DB 1 aatkqegmglitglvqlgsddinnvtcaagilsnltc 39

RESULT 2
 AAY33236
 ID AAY33236 standard; peptide; 39 AA.
 XX
 AC AAY33236;
 XX
 DT 18-NOV-1999 (first entry)
 XX
 DE Human beta-catenin protein mutant armadillo repeat arm 7.
 XX
 KW Beta-catenin; human; armadillo repeat; treatment; human disease; LEF-1;
 KW modulator; interaction domain; transcription factor; TCF-4; oncogenic;
 KW tumor suppressor gene; conductin; E-cadherin; screening; carcinoma; APC;
 KW Wnt signalling pathway; tumor development; anti-oncogenic; melanoma;
 KW organ regeneration; tissue regeneration; hair growth; mutant.
 XX
 OS Homo sapiens.
 OS Synthetic.
 PN DE19909251-A1.
 XX
 PD 26-AUG-1999.
 XX
 PF 22-FEB-1999; 99DE-1009251.
 XX
 PR 21-FEB-1998; 98DE-1007390.
 XX
 PA (DELB-) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.
 XX
 PI Birchmeier W, Von Kries J;
 DR WPI; 1999-470389/40.
 XX
 DR Agents for treating human diseases, particularly cancer, modulate
 PT interaction of beta-catenin with transcription factors or tumor

PT suppressor gene products
 XX Example 4; Fig 5; 16pp; German.
 PS This invention describes a novel agent (A) for treating human disease
 XX which is based on substances (I) that modulate (inhibit or promote) the
 CC interaction of beta-catenin with transcription factors or products of
 CC tumor suppressor genes. The invention also describes (a) peptides (II)
 CC comprising part of the LEF-1/TCF-4 transcription factors, or their
 CC variants and mutants; (b) peptides and related molecules (III) from the
 CC armadillo domain (arm units 3-8) of beta-catenin, and mutants of the
 CC complete beta-catenin molecule, that include at least one of the
 CC specific interaction domains for LEF-1, TCF-4, APC, conductin or
 CC E-cadherin; and (c) enzyme-linked immunosorbent assay (ELISA) for
 CC screening substance libraries for compounds that modulate interaction of
 CC beta-catenin with LEF-1/TCF, APC, conductin or E-cadherin. Beta-catenin
 CC is a key compound in the Wnt signalling pathway and is involved in
 CC development of tumors. Generally its interaction with LEF-1 or TCF-4 is
 CC oncogenic but interaction with APC, conductin or E-cadherin is
 CC anti-oncogenic. (A) which inhibit interaction are particularly used to
 CC treat tumors, especially carcinoma of the colon and melanoma, but also,
 CC where they promote interaction, to stimulate regeneration of organs and
 CC tissues, specifically hair growth. AAY33230-Y33241 represent mutant
 CC human beta-catenin armadillo repeat fragments described in the method of
 CC the invention.
 XX
 SQ Sequence 39 AA;

Query Match 100.0%; Score 194; DB 20; Length 39;
 Best Local Similarity 100.0%; Pred. No. 4.7e-21;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATKQEGMGLLTGLVQLGSDDDINVTCAAGILSNLTC 39
 |||||
 DB 1 aatkqegmglitglvqlgsddinnvtcaagilsnltc 39

RESULT 3
 AAB07290
 ID AAB07290 standard; Protein; 781 AA.
 XX
 AC AAB07290;
 XX
 DT 01-NOV-2000 (first entry)
 XX
 DE Human beta catenin.
 XX
 KW Beta catenin; cadherin; metastasis; cadherin-associated protein;
 KW human; colorectal cancer; melanoma; antisense oligonucleotide;
 KW gene therapy.
 XX
 OS Homo sapiens.
 XX US6066500-A.
 PN 23-MAY-2000.
 XX
 PD 25-JUN-1999; 99US-0344519.
 XX
 PR 25-JUN-1999; 99US-0344519.
 XX
 PA (ISIS-) ISIS PHARM INC;
 XX
 PI Bennett CF, Cowser LM;
 XX
 DR WPI; 2000-410651/35;
 DR N-PSDB; AAA58320.
 XX
 XX New antisense compounds targeting nucleic acids encoding human beta
 PT catenin (HBC) useful for treating diseases associated with HBC
 PT expression and as prophylaxis to prevent or delay infection,
 PT inflammation or tumor formation

XX

PS Example 13; Columns 45-52; 35pp; English.

XX

Beta catenin is a member of the catenin family of cytosolic proteins and a key member of the Wnt signalling pathway. Catenins interact with the cytoplasmic domains of cadherin glycoproteins, and are important in maintaining cell adhesiveness. The loss of cell adhesiveness is implicated in metastasis. Beta catenin is also known as cadherin-associated protein and is implicated in colorectal cancer and melanoma. The present sequence is the human beta catenin protein. The coding sequence of this protein was used in the present invention to design antisense oligonucleotides (AA58327-AS8366). The oligonucleotides are capable of hybridising to human beta catenin, in order to inhibit expression of human beta catenin. The oligonucleotides may be used in gene therapy for colorectal cancer or melanoma.

XX

SQ Sequence 781 AA;

Query Match

Best Local Similarity 100.0%; Score 194; DB 21; Length 781;

Matches

39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 AATKQEGMGLLGLTIVQLGSDDDINVTCAAGILSNLTC 39

Db

391 aatkqegmglgtivqlgsgddinvtcaagilsnltc 429

RESULT 4

AA70740

ID

AA70740 standard; protein; 781 AA.

XX

AC AA70740;

XX

DT 24-JUL-2000 (first entry)

XX

DE Human beta-catenin.

XX

Human beta-catenin; Wnt antagonist; contraceptive; contraceptive vaccine; oocyte development; female primate contraception; oocyte viability; monoclonal antibody; Wnt signalling.

XX

OS Homo sapiens.

XX

PN WO200021555-A1.

XX

PD 20-APR-2000.

XX

PF 13-OCT-1999; 99WO-US23640.

XX

PR 15-OCT-1998; 98US-0104355.

XX

PA (HARD) HARVARD COLLEGE.

XX

PI McMahon AP, Parr BA, Vaino S;

XX

DR WPI; 2000-317845/27.

XX

Contraceptive composition for inhibiting oocyte development in a female primate comprises a Wnt polypeptide antagonist -

XX

Example 3; Page 26; 57pp; English.

XX

The patent discloses a method of female primate contraception comprising administering an antagonist of a Wnt polypeptide, inhibiting oocyte development. Wnt polypeptides are useful for promotive maturation of an immature oocyte. Wnt polypeptides are also useful for increasing the number of mature oocytes and to enhance oocyte viability. Soluble fragments of Wnt polypeptides have the ability to inhibit Wnt signalling, e.g., by blocking binding of a naturally-occurring Wnt protein to its receptor. They may be used to generate monoclonal antibodies which can inhibit oocyte development. The present sequence is the human beta-catenin protein. Dominant negative mutants of beta-catenin lack one

CC

or more armidillo-like repeats which participate in cadherin binding. Other mutants include those lacking amino acids 555-781 or 424-781, or 1-422 of beta-catenin.

CC

SQ Sequence 781 AA;

Query Match

Best Local Similarity 100.0%; Score 194; DB 21; Length 781;

Matches

39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 AATKQEGMGLLGLTIVQLGSDDDINVTCAAGILSNLTC 39

Db

391 aatkqegmglgtivqlgsgddinvtcaagilsnltc 429

RESULT 5

AAU28118

ID

AAU28118 standard; Protein; 781 AA.

XX

AC AAU28118;

XX

DT 18-DEC-2001 (first entry)

XX

DE Novel human secretory protein, Seq ID No 287.

XX

Human; secreted protein; arthritis; Crohn's disease; sepsis; shock; ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy; transgenic animal; Alzheimer's disease; Parkinson's disease; burn; amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia; ulcer; osteoporosis; bone degenerative disorder; periodontal disease; gut protection; lung; liver fibrosis; immune deficiency; infection; severe combined immunodeficiency; SCID; autoimmune disorder; allergy; multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma; fertility; analgesic; pain; antigen.

XX

OS Homo sapiens.

XX

PN WQ200166689-A2.

XX

PD 13-SEP-2001.

XX

PF 05-MAR-2001; 2001WO-US04942.

XX

PR 07-MAR-2000; 2000US-0519705.

XX

PR 19-MAY-2000; 2000US-0574454.

XX

PR 17-JUN-2000; 2000US-0596193.

XX

PR 14-JUL-2000; 2000US-0616847.

XX

PR 19-SEP-2000; 2000US-0665363.

XX

PR 20-OCT-2000; 2000US-0693267.

XX

PA (HYSE-) HYSEQ INC.

XX

Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;

XX

Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;

XX

WPI; 2001-589934/66.

XX

N-PSDB; AA545018.

XX

Novel polypeptides and nucleic acids obtained from cDNA libraries prepared from various human tissues, for diagnosis and treatment of cancer, neurological, inflammatory, and autoimmune disorders -

XX

Example 4; SEQ ID No 287; 107pp; English.

XX

The invention relates to novel isolated human secreted polypeptides (I) and polynucleotides (II). (I) and (II) are useful for treating inflammatory conditions such as arthritis, nephritis, Crohn's disease, ischaemia-reperfusion injury, shock, sepsis, immune responses, and is involved in increasing haematopoiesis, stem cell survival, bone growth and remodeling. (I), (II) and modulators of (II) are useful for prophylaxis or treatment of one or more cancers. (II) is also useful for creating transgenic animals useful for studying the in vivo activities of

the polypeptide as well as for studying modulators of the polypeptides.
 (1) induces the proliferation of neural cells and regeneration of nerve
 and brain tissue and is useful for the treatment of central and
 peripheral nervous system diseases and neuropathies, such as Alzheimer's,
 Parkinson's disease, Huntington's disease, and amyotrophic lateral
 sclerosis. In addition, (1) is involved in chemotactic or chemokinetic
 activity, regulation of haematopoiesis and is useful for treating myeloid
 or lymphoid cell disorders, platelet disorders such as thrombocytopenia
 and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 tissue growth, and in tissue repair, healing of burns, incisions,
 ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
 disorders, or periodontal disease. Furthermore, (1) is also useful for
 gut protection or regeneration and treatment of lung or liver fibrosis,
 reperfusion injury in various tissues, various immune deficiencies and
 disorders including severe combined immunodeficiency (SCID), bacterial or
 fungal infections, autoimmune disorders e.g. multiple sclerosis,
 rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
 reactions and conditions, such as asthma or other respiratory problems.
 In addition, (1) affects biorhythms or circadian cycles of rhythms,
 fertility, metabolism, catabolism, anabolism, storage or elimination of
 dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
 analgesic effects or other pain reducing effects, immunoglobulin like
 activity and can act as an antigen in a vaccine composition to raise an
 immune response. AAU28020-AAU28395 represent novel human secreted protein
 amino acid sequences of the invention.

XX Sequence 781 AA;

Query Match 100.0%; Score 194; DB 22; Length 781;
 Best Local Similarity 100.0%; Pred. No. 1.6e-19;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATKQEGMEGLLTGLVQLGSDDDINVTCAAGILSLNLT 39
 |||||
 DB 391 aatkqegmglitglvqlgsddinvtcaagilslntc 429

RESULT 6
 AAEO6038
 ID AAE06038 standard; Protein; 781 AA.

XX AC AAE06038;

XX DT 25-SEP-2001 (first entry)

XX DE Human beta-catenin protein #1.

XX Human; stem cell culturing; progenitor cell; pluripotential phenotype;
 KW transplantation; haematopoietic function; allogeneic recipient;
 KW signalling pathway; beta-catenin.

XX OS Homo sapiens.

XX PN WO200152649-A1.

XX PD 26-JUL-2001.

XX PF 17-JAN-2001; 2001WO-US01459.

XX PR 18-JAN-2000; 2000US-0176786.

XX PA (STRD) UNIV LELAND STANFORD JUNIOR.

XX PI Reya T, Nusse R, Weissman IL;

XX WPI: 2001-465328/50.

DR N-PSDB; AAD11164.

XX In vitro expansion of mammalian stem or progenitor cells, useful for
 PT producing cells that retain their pluripotential phenotype after
 PT expansion for use in transplantation, involves increasing the levels of
 PT beta-catenin in the cell -

XX

PS Disclosure; Page 23-24; 33pp; English.

XX

CC The present invention relates to a method for in vitro expansion of
 CC mammalian stem or progenitor cells, which comprising increasing the
 CC intracellular concentration of beta-catenin in a progenitor or stem cell
 CC in an in vitro culture medium for a period sufficient for the progenitor
 CC or stem cell to divide. The number of cells having the functional
 CC phenotype of the stem or progenitor cells is expanded. The method is
 CC useful for culturing stem cells and progenitor cells, which retain their
 CC pluripotential phenotype after expansion, in vitro. The expanded cell
 CC population are useful as a source of stem cells, e.g. to reconstitute
 CC function in a host that is deficient in a particular cell lineage or
 CC lineages. The expanded cell populations are also useful in
 CC transplantation to restore haematopoietic function to autologous or
 CC allogeneic recipients. The present sequence is human beta-catenin
 CC protein. Beta-catenin is a pivotal player in the signalling pathway
 CC initiated by Wnt proteins, which are mediators of several developmental
 CC processes.

XX Sequence 781 AA;

Query Match 100.0%; Score 194; DB 22; Length 781;

Best Local Similarity 100.0%; Pred. No. 1.6e-19;

Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATKQEGMEGLLTGLVQLGSDDDINVTCAAGILSLNLT 39

|||||

DB 391 aatkqegmglitglvqlgsddinvtcaagilslntc 429

RESULT 7

AAEO6039

ID AAE06039 standard; Protein; 781 AA.

XX AC AAE06039;

XX DT 25-SEP-2001 (first entry)

XX DE Mouse beta-catenin protein #2.

XX Mouse; stem cell culturing; progenitor cell; pluripotential phenotype;

KW transplantation; haematopoietic function; allogeneic recipient;

KW signalling pathway; beta-catenin.

XX OS Mus musculus.

XX PN WO200152649-A1.

XX PD 26-JUL-2001.

XX PF 17-JAN-2001; 2001WO-US01459.

XX PR 18-JAN-2000; 2000US-0176786.

XX PA (STRD) UNIV LELAND STANFORD JUNIOR.

XX PI Reya T, Nusse R, Weissman IL;

XX WPI: 2001-465328/50.

DR N-PSDB; AAD11165.

XX In vitro expansion of mammalian stem or progenitor cells, useful for
 PT producing cells that retain their pluripotential phenotype after
 PT expansion for use in transplantation, involves increasing the levels of
 PT beta-catenin in the cell -

PS Disclosure; Page 28-30; 33pp; English.

XX The present invention relates to a method for in vitro expansion of

CC mammalian stem or progenitor cells, which comprising increasing the

CC intracellular concentration of beta-catenin in a progenitor or stem cell

CC in an in vitro culture medium for a period sufficient for the progenitor
 CC or stem cell to divide. The number of cells having the functional
 CC phenotype of the stem or progenitor cells is expanded. The method is
 CC useful for culturing stem cells and progenitor cells, which retain their
 CC pluripotential phenotype after expansion, in vitro. The expanded cell
 CC populations are useful as a source of stem cells, e.g. to reconstitute
 CC function in a host that is deficient in a particular cell lineage or
 CC lineages. The expanded cell populations are also useful in
 CC transplantation to restore haematopoietic function to autologous or
 CC allogeneic recipients. The present sequence is mouse beta-catenin
 CC protein. Beta-catenin is a pivotal player in the signalling pathway
 CC initiated by Wnt proteins, which are mediators of several developmental
 CC processes.

XX Sequence 781 AA;

Query Match 100.0%; Score 194; DB 22; Length 781;
 Best Local Similarity 100.0%; Pred. No. 1.6e-19;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATKQEGMEGLLTGLVQLLGSDDINVTCAAGILSNLTC 39

Db 391 aatkqegmeglgtlvtllgsddinvvtcaagilslntc 429

RESULT 8

AAU28306
 ID AAU28306 standard; Protein: 800 AA.

XX AC AAU28306;

XX 18-DEC-2001 (first entry)

XX Novel human secretory protein, Seq ID No 663.

XX Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
 KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
 KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
 KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
 KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
 KW gut protection; lung; liver fibrosis; immune deficiency; infection;
 KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
 KW fertility; analgesic; pain; antigen.

XX Homo sapiens.

XX WO200166689-A2.

XX 13-SEP-2001.

XX 05-MAR-2001; 2001WO-US04942.

XX 07-MAR-2000; 2000US-0519705.

PR 19-MAY-2000; 2000US-0574454.

PR 17-JUN-2000; 2000US-0596193.

PR 14-JUL-2000; 2000US-0616847.

PR 19-SEP-2000; 2000US-0665363.

PR 20-OCT-2000; 2000US-0693267.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;

PI Zhao QN, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;

XX WPI; 2001-589934/66.

DR N-PSDB; AAS45206.

XX Novel polypeptides and nucleic acids obtained from cDNA libraries
 PT prepared from various human tissues, for diagnosis and treatment of
 PT cancer, neurological, inflammatory, and autoimmune disorders -

XX

Example 2; SEQ ID NO 663; 107pp; English.

PS The invention relates to novel isolated human secreted polypeptides (I)
 XX and polynucleotides (II). (I) and (II) are useful for treating
 CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
 CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
 CC involved in increasing haematopoiesis, stem cell survival, bone growth
 CC and remodelling. (I), (II) and modulators of (II) are useful for
 CC prophylaxis or treatment of one or more cancers. (II) is also useful for
 CC creating transgenic animals useful for studying the in vivo activities of
 CC the polypeptide as well as for studying modulators of the polypeptides.
 CC (I) induces the proliferation of neural cells and regeneration of nerve
 CC and brain tissue and is useful for the treatment of central and
 CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
 CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
 CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
 CC activity, regulation of haematopoiesis and is useful for treating myeloid
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 CC tissue growth, and in tissue repair, healing of burns, incisions,
 CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
 CC disorders, or periodontal disease. Furthermore, (I) is also useful for
 CC gut protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues, various immune deficiencies and
 CC disorders including severe combined immunodeficiency (SCID), bacterial or
 CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
 CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
 CC reactions and conditions, such as asthma or other respiratory problems.
 CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
 CC fertility, metabolism, catabolism, anabolism, storage or elimination of
 CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
 CC analgesic effects or other pain reducing effects, immunoglobulin like
 CC activity and can act as an antigen in a vaccine composition to raise an
 CC immune response. AAU28020-AAU28395 represent novel human secreted protein
 CC amino acid sequences of the invention.

XX Sequence 800 AA;

Query Match 100.0%; Score 194; DB 22; Length 800;
 Best Local Similarity 100.0%; Pred. No. 1.6e-19;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATKQEGMEGLLTGLVQLLGSDDINVTCAAGILSNLTC 39

Db 404 aatkqegmeglgtlvtllgsddinvvtcaagilslntc 442

RESULT 9

ABB60196

ID AB560196 standard; Protein: 840 AA.

XX AC ABB60196;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 7380.

DE Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE?) PE CORP NY.

PA

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 16:05:10 ; Search time 57.6 Seconds
(without alignments)
16.538 Million cell updates/sec

Title: US-09-641-104A-10
Perfect score: 194
Sequence: 1 AATKQEGMEGLTGLVQLGSDINVTCAAGILSNLTC 39

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pgp.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pgp.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pgp.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pgp.*
5: /cgn2_6/ptodata/2/1aa/PT05_COMB.pgp.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pgp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	52	26.8	11	4	US-09-296-089-28	Sequence 28, Appl
2	51	26.3	178	2	US-08-630-822A-100	Sequence 100, App
3	51	26.3	178	2	US-09-005-069-100	Sequence 100, App
4	50	25.8	529	2	US-08-933-227-4	Sequence 4, Appli
5	48.5	25.0	1012	1	US-08-219-262B-3	Sequence 3, Appli
6	48.5	25.0	1012	1	US-08-219-262B-5	Sequence 5, Appli
7	48.5	25.0	1012	2	US-08-708-541A-30	Sequence 30, Appl
8	48.5	25.0	1012	3	US-09-031-655-3	Sequence 3, Appli
9	48.5	25.0	1012	3	US-09-031-655-5	Sequence 5, Appli
10	48	24.7	398	3	US-09-189-035-6	Sequence 6, Appli
11	48	24.7	398	4	US-09-382-086-6	Sequence 6, Appli
12	47.5	24.5	483	4	US-09-108-020-16	Sequence 16, Appl
13	47.5	24.5	666	4	US-08-982-785A-11	Sequence 11, Appl
14	47.5	24.5	1012	3	US-08-219-262B-7	Sequence 7, Appli
15	47.5	24.5	1012	3	US-09-031-655-7	Sequence 7, Appli
16	47	24.2	837	2	US-08-216-260-2	Sequence 2, Appli
17	47	24.2	2842	1	US-07-741-940-7	Sequence 7, Appli
18	47	24.2	2842	1	US-08-289-548A-7	Sequence 7, Appli
19	47	24.2	2842	1	US-08-452-654-7	Sequence 7, Appli
20	47	24.2	2843	1	US-07-741-940-2	Sequence 2, Appli
21	47	24.2	2843	1	US-08-289-548A-2	Sequence 2, Appli
22	47	24.2	2843	1	US-08-452-654-2	Sequence 2, Appli
23	47	24.2	2843	1	US-08-452-655B-2	Sequence 2, Appli
24	47	24.2	2843	1	US-08-452-655B-7	Sequence 7, Appli
25	47	24.2	2843	3	US-08-370-235A-2	Sequence 2, Appli
26	47	24.2	2843	3	US-08-450-582-2	Sequence 2, Appli
27	47	24.2	2843	3	US-08-450-582-7	Sequence 7, Appli

28 47 24.2 2973 2 US-08-821-355A-7 Sequence 7, Appli
29 47 24.2 2973 2 US-09-003-687A-7 Sequence 7, Appli
30 47 24.2 2973 4 US-09-136-605-7 Sequence 7, Appli
31 46.5 24.0 3119 1 US-08-246-982A-16 Sequence 16, Appli
32 46.5 24.0 3119 1 US-08-453-263-16 Sequence 16, Appli
33 46 23.7 379 4 US-09-740-035-4 Sequence 4, Appli
34 46 23.7 989 4 US-09-110-517-4 Sequence 4, Appli
35 46 23.7 1281 2 US-08-843-530B-6 Sequence 6, Appli
36 46 23.7 1298 2 US-08-843-530B-2 Sequence 2, Appli
37 46 23.7 1298 2 US-08-843-530B-4 Sequence 4, Appli
38 46 23.7 1298 2 US-08-843-530B-34 Sequence 34, Appli
39 46 23.7 2020 1 US-07-551-531-2 Sequence 1, Appli
40 46 23.7 2485 5 PCT-US94-00198-1 Sequence 1, Appli
41 46 23.7 2485 5 PCT-US94-00198-2 Sequence 2, Appli
42 46 23.7 2818 1 US-08-510-284-1 Sequence 1, Appli
43 46 23.7 2818 1 US-08-411-389-2 Sequence 2, Appli
44 46 23.7 2818 2 US-08-449-933-2 Sequence 2, Appli
45 46 23.7 2818 4 US-07-966-049A-2 Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-296-089-28
; Sequence 28, Application US/09296089
; Patent No. 6303576
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Byers, Stephen
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
; TITLE OF INVENTION: BETA-CATENIN MEDIATED GENE EXPRESSION
; FILE REFERENCE: 100086.411
; CURRENT APPLICATION NUMBER: US/09/296,089
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 28
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-296-089-28

Query Match 26.8%; Score 52; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 12 LGTLVOLLGSD 22
Db 1 LGTLVOLLGSD 11.
RESULT 2
US-08-630-822A-100
; Sequence 100, Application US/08630822A
; Patent No. 5840695
; GENERAL INFORMATION:
; APPLICANT: HUNTER, SHIRLEY WU
; APPLICANT: WALLENFELS, LYNDIA
; TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 11-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CONNELL, GARY J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-17-C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 100:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 178 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-630-822A-100

```

Query Match 26.3%; Score 51; DB 2; Length 178;
 Best Local Similarity 40.0%; Pred. No. 4.9;
 Matches 14; Conservative 7; Mismatches 10; Indels 4; Gaps 1;

```

QY 8 MEGGLGLTVQLL---GSDINNVVTCAGILSNLT 38
   ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5 VEGALGFLVEMLYKAPSKTLAIVENAGGILRNVS 39

```

```

RESULT 3
US-09-005-069-100
; Sequence 100, Application US/090005069
; Patent No. 5932470
; GENERAL INFORMATION:
; APPLICANT: FRANK, GLENN R.
; APPLICANT: HUNTER, SHIRLEY WU
; APPLICANT: WALLENFELS, LYNDIA
; TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS
; TITLE OF INVENTION: AND APPARATUS TO COLLECT SUCH PROTEINS
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/005,069
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/630,822
; FILING DATE: 11-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: CONNELL, GARY J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-17-C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 100:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 178 amino acids

```

```

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-005-069-100

```

Query Match 26.3%; Score 51; DB 2; Length 178;
 Best Local Similarity 40.0%; Pred. No. 4.9;
 Matches 14; Conservative 7; Mismatches 10; Indels 4; Gaps 1;

```

QY 8 MEGGLGLTVQLL---GSDINNVVTCAGILSNLT 38
   ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5 VEGALGFLVEMLYKAPSKTLAIVENAGGILRNVS 39

```

```

RESULT 4
US-08-933-227-4
; Sequence 4, Application US/08933227
; Patent No. 5965394
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Guegler, Karl
; APPLICANT: Corley, Neil
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN IMPORTIN ALPHA HOMOLOG
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,227
; FILING DATE: Filed Herewith
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0394 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 529 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1708480
; US-08-933-227-4

```

Query Match 25.8%; Score 50; DB 2; Length 529;
 Best Local Similarity 39.4%; Pred. No. 26;
 Matches 13; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

```

QY 6 EGMEGLGLTVQLLGGSDINNVVTCAGILSNLT 38
   :|||: |||: |||: |||: |||: |||: |||: |||
Db 247 DAVEQILPTLVRLHDDPEVLADTCWAISYLT 279

```

RESULT 5

US-08-219-262B-3
; Sequence 3, Application US/08219262B
; Patent No. 5788970
; GENERAL INFORMATION:
; APPLICANT: VAKHARIA, VIKRAM
; APPLICANT: SNYDER, DAVID B
; APPLICANT: MENDEL-WHERSAT, STEPHANIE A
; TITLE OF INVENTION: CHIMERIC INFECTIOUS BURSAL DISEASE VIRUS
; TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED
; TITLE OF INVENTION: THEREON
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
; STREET: 1755 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/219,262B
; FILING DATE: 29-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 2747-047-27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1012 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Infectious bursal disease virus
; STRAIN: E/DEL
US-08-219-262B-3

Query Match 25.0%; Score 48.5; DB 1; Length 1012;
Best Local Similarity 37.1%; Pred. No. 98;
Matches 13; Conservative 5; Mismatches 16; Indels 1; Gaps 1;

QY 5 QEGMEGL-LGTLVQLGSDDDINVTCAAGILSNLT 38
| : | | | : | | | : | | : | |
Db 249 QTSVHGLVGLGATYILIGFDGTTVITRAVAANNGLT 283

RESULT 6
US-08-219-262B-5
; Sequence 5, Application US/08219262B
; Patent No. 5788970
; GENERAL INFORMATION:
; APPLICANT: VAKHARIA, VIKRAM
; APPLICANT: SNYDER, DAVID B
; APPLICANT: MENDEL-WHERSAT, STEPHANIE A
; TITLE OF INVENTION: CHIMERIC INFECTIOUS BURSAL DISEASE VIRUS
; TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED
; TITLE OF INVENTION: THEREON
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
; STREET: 1755 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR
; CITY: ARLINGTON

US-08-219-262B-3
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/219,262B
; FILING DATE: 29-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 2747-047-27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1012 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Infectious bursal disease virus
; STRAIN: CU-1
US-08-219-262B-5

Query Match 25.0%; Score 48.5; DB 1; Length 1012;
Best Local Similarity 37.1%; Pred. No. 98;
Matches 13; Conservative 5; Mismatches 16; Indels 1; Gaps 1;

QY 5 QEGMEGL-LGTLVQLGSDDDINVTCAAGILSNLT 38
| : | | | : | | | : | | : | |
Db 249 QTSVHGLVGLGATYILIGFDGTTVITRAVAANNGLT 283

RESULT 7
US-08-708-541A-30
; Sequence 30, Application US/08708541A
; Patent No. 5871744
; GENERAL INFORMATION:
; APPLICANT: VAKHARIA, Vikram N.
; APPLICANT: MUNDT, Egbert
; TITLE OF INVENTION: A METHOD FOR GENERATING BIRNAVIRUS FROM
; TITLE OF INVENTION: SYNTHETIC RNA TRANSCRIPTS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIKAIKO, MARCELSTEIN, MURRAY & ORAM LLP
; STREET: 655 Fifteenth Street, N. W.,
; STREET: Suite 330 - G Street Lobby
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/708,541A
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: KITTS, Monica C.
; REGISTRATION NUMBER: 36,105
; REFERENCE/DOCKET NUMBER: .P8172-6002

us-09-641-104a-10.ra1

Tue Jul 30 08:15:22 2002

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202/638-5000
TELEFAX: 202/638-4810
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 1012 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-708-541A-30

Query Match
Best Local Similarity: 37.1%; Pred. No. 98;
Matches: 13; Conservative 5; Mismatches 16; Indels 1; Gaps 1;

QY 5 QEGMEGL-LGTLVOLLGSDDDINNVTCAGILSNLT 38
DB 249 QTSVHGLVGLGATYILIGFDGTTVITRAVAANGLT 283

RESULT: 9
US-09-031-655-5
Sequence 5, Application US/09031655
Patent No. 6017759
GENERAL INFORMATION:
APPLICANT: VAKHARIA, VIKRAM
APPLICANT: MENDEL-WHERRAS, STEPHANIE A
TITLE OF INVENTION: CHIMERIC INFECTIOUS BURSAL DISEASE VIRUS
TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
STREET: 1755 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: USA
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/031,655
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/219,262
FILING DATE: 29-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 2747-047-27
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1012 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Infectious bursal disease virus
STRAIN: CU-1
US-09-031-655-5

Query Match
Best Local Similarity: 37.1%; Pred. No. 98;
Matches: 13; Conservative 5; Mismatches 16; Indels 1; Gaps 1;

QY 5 QEGMEGL-LGTLVOLLGSDDDINNVTCAGILSNLT 38
DB 249 QTSVHGLVGLGATYILIGFDGTTVITRAVAANGLT 283

RESULT: 10
US-09-189-035-6
Sequence 6, Application US/09189035

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202/638-5000
TELEFAX: 202/638-4810
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 1012 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-708-541A-30

Query Match
Best Local Similarity: 37.1%; Pred. No. 98;
Matches: 13; Conservative 5; Mismatches 16; Indels 1; Gaps 1;

QY 5 QEGMEGL-LGTLVOLLGSDDDINNVTCAGILSNLT 38
DB 249 QTSVHGLVGLGATYILIGFDGTTVITRAVAANGLT 283

RESULT: 10
US-09-189-035-6
Sequence 6, Application US/09189035

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202/638-5000
TELEFAX: 202/638-4810
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 1012 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-708-541A-30

Query Match
Best Local Similarity: 37.1%; Pred. No. 98;
Matches: 13; Conservative 5; Mismatches 16; Indels 1; Gaps 1;

QY 5 QEGMEGL-LGTLVOLLGSDDDINNVTCAGILSNLT 38
DB 249 QTSVHGLVGLGATYILIGFDGTTVITRAVAANGLT 283

RESULT: 10
US-09-189-035-6
Sequence 6, Application US/09189035

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202/638-5000
TELEFAX: 202/638-4810
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 1012 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-708-541A-30

Query Match
Best Local Similarity: 37.1%; Pred. No. 98;
Matches: 13; Conservative 5; Mismatches 16; Indels 1; Gaps 1;

QY 5 QEGMEGL-LGTLVOLLGSDDDINNVTCAGILSNLT 38
DB 249 QTSVHGLVGLGATYILIGFDGTTVITRAVAANGLT 283

RESULT: 10
US-09-189-035-6
Sequence 6, Application US/09189035

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202/638-5000
TELEFAX: 202/638-4810
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 1012 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-708-541A-30

Query Match
Best Local Similarity: 37.1%; Pred. No. 98;
Matches: 13; Conservative 5; Mismatches 16; Indels 1; Gaps 1;

QY 5 QEGMEGL-LGTLVOLLGSDDDINNVTCAGILSNLT 38
DB 249 QTSVHGLVGLGATYILIGFDGTTVITRAVAANGLT 283

RESULT: 10
US-09-189-035-6
Sequence 6, Application US/09189035

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202/638-5000
TELEFAX: 202/638-4810
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 1012 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-708-541A-30

Query Match
Best Local Similarity: 37.1%; Pred. No. 98;
Matches: 13; Conservative 5; Mismatches 16; Indels 1; Gaps 1;

RESULT 12
US-09-108-020-16
; Sequence 16, Application US/09108020A
; Patent No. 6143561
; GENERAL INFORMATION:
; APPLICANT: Randall, Douglas D.
; APPLICANT: Johnston, Mark L.
; APPLICANT: Miernyk, Jan A.

QY 11 LLGTLVQLLGSDDINV-----VTCAAGILSNLTC 39
||| | : | :: | : ||| : | |
Dd 20 LLGLRTGVAGEDLOVIOPEKSVVAAGESATIRC 54

NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FASTSQ for Windows version 2.0 b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/982.785A
FILING DATE: 02-DEC-1997
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER: US 60/031,556
FILING DATE: 02-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 05311/018001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
TOPOLOGY: linear

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 16:09:14 ; Search time 70.63 Seconds
(without alignments)
53.058 Million cell updates/sec

Title: US-09-641-104a-10
Perfect score: 194
Sequence: 1 AATKQEGMEGLTGLVQLLGSDDDINVTCAAGILSNLTC 39
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	194	100.0	781	2 S35099	beta-catenin - Afr
2	194	100.0	781	2 A38973	beta-catenin - hum
3	194	100.0	781	2 S35091	beta-catenin - mou
4	155	79.9	843	2 T12689	armadillo segment
5	150	77.3	806	2 JC4835	beta-catenin - Hyd
6	147	75.8	817	2 S33793	hypothetical prote
7	131	67.5	738	2 S35093	plakoglobin - Afri
8	128	66.0	744	2 A32905	plakoglobin, desmo
9	120	61.9	820	2 S33794	hypothetical prote
10	113	58.2	621	2 S35092	plakoglobin - mous
11	71	36.6	678	2 T23341	beta-catenin - Cae
12	64	33.0	796	2 T03746	hypothetical prote
13	64	33.0	964	2 T15342	hypothetical prote
14	62	32.0	1299	2 A86366	T26J12.6 protein -
15	56.5	29.1	783	2 A81961	zam protein (impor
16	56	28.9	322	2 S77022	hypothetical prote
17	55.5	28.6	248	2 B87351	hypothetical prote
18	55	28.4	526	2 AD2042	phosphoglycerate d
19	55	28.4	555	2 T40294	hypothetical prote
20	54.5	28.1	868	2 S76777	aconitate hydratase
21	54	27.8	419	2 T39248	probable GTP bindi
22	54	27.8	458	2 T04573	hypothetical prote
23	54	27.8	525	1 C69705	phosphoglycerate d
24	53.5	27.6	526	2 F84385	dipeptide ABC tran
25	53	27.3	166	2 B69886	conserved hypothet
26	53	27.3	253	2 AG1346	dehydrogenase homo
27	53	27.3	554	2 S75016	phosphoglycerate d
28	52.5	27.1	538	2 AG3295	D-3-phosphoglycer
29	52.5	27.1	983	2 A87063	conserved large me

30	52.5	27.1	1008	2 S72698	transport protein
31	52	26.8	238	2 E64416	hypothetical prote
32	52	26.8	556	2 T47552	hypothetical prote
33	52	26.8	664	2 T50316	hypothetical Armad
34	51.5	26.5	487	2 I64033	hypothetical prote
35	51.5	26.5	514	2 B90097	importin alpha (im
36	51.5	26.5	578	2 A34283	H+-transporting Ar
37	51.5	26.5	579	2 A99093	subtilisinlike seri
38	51	26.3	338	2 T36307	hypothetical prote
39	51	26.3	885	2 D86151	F22M8.8 protein -
40	51	26.3	2207	2 T42759	Munc13-3 protein -
41	50.5	26.0	80	2 S59961	antimicrobial pept
42	50	25.8	133	2 C83100	hypothetical prote
43	50	25.8	214	2 AH0265	probable membrane
44	50	25.8	215	2 AF0650	probable membrane
45	50	25.8	295	2 S60711	band-6-protein - b

ALIGNMENTS

RESULT 1

S35099
beta-catenin - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C>Date: 10-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 21-Jul-2000
C:Accession: S35099
R:McCreary P.D.; Turck, C.W.; Gumbiner, B.
Science 254, 1359-1361, 1991
A:Title: A homolog of the armadillo protein in Drosophila (plakoglobin) associated w/
A:Reference number: S35099; MUID:92073903
A:Accession: S35099
A:Molecule type: mRNA
A:Residues: 1-781 <NCC>
A:Cross-references: GB:M77013; NID:g214020; PIDN:AAA49670.1; PID:g214021
C:Keywords: cytoskeleton

Query Match 100.0%; Score 194; DB 2; Length 781;
Best Local Similarity 100.0%; Pred. No. 8.2e-18;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATKQEGMEGLTGLVQLLGSDDDINVTCAAGILSNLTC 39
|||||
Db 391 AATKQEGMEGLTGLVQLLGSDDDINVTCAAGILSNLTC 429

RESULT 2

A38973
beta-catenin - human
C:Species: Homo sapiens (man)
C>Date: 26-Jan-1996 #sequence_revision 08-Feb-1996 #text_change 08-Oct-1999
C:Accession: A38973; S55356; S31988
R:Huelsken, J.; Birchmeier, W.; Behrens, J.
J. Cell Biol. 127, 2061-2069, 1994
A:Title: E-cadherin and APC compete for the interaction with beta-catenin and the cyt
A:Reference number: A38973; MUID:95105247
A:Accession: A38973
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-781 <HUE>
A:Cross-references: GB:219054; NID:g38519; PIDN:CAA79497.1; PID:g38520
A:Experimental source: placenta
R:Nollet, F.; Berx, G.; Molemans, F.; van Roy, F.
submitted to the EMBL Data Library, June 1995
A:Description: H. sapiens beta-catenin mRNA.
A:Reference number: S55356
A:Accession: S55356
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-781 <NOL>
A:Cross-references: EMBL:X87838; NID:g1154853; PIDN:CAA61107.1; PID:g860988
C:Comment: The cytoplasmic protein beta-catenin normally binds to a region in the cyt

microfilament network.
C:Comment: Cellular levels of beta-catenin are regulated in part by the adenomatous poly
d cancerous cell growth.

C:Genetics:
A:Gene: GDB:CTNNB1; CTNNB
A:Cross-references: GDB:141922; OMIM:116806
A:Map position: 3p22-3p21.3
C:Keywords: apoptosis; carcinogenesis; cell adhesion; cytosol
F:151-676/Region: 40-residue repeats

Query Match 100.0%; Score 194; DB 2; Length 781;
Best Local Similarity 100.0%; Pred. No. 8.2e-18;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATKQEGMEGLGLTVQLVLLGSDDDINVTCAAGILSNLTC 39
|||||
Db 391 AATKQEGMEGLGLTVQLVLLGSDDDINVTCAAGILSNLTC 429

RESULT 3
beta-catenin - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999
C:Accession: S35091
R:Butz, S.; Stappert, J.; Weissig, H.; Kemler, R.
Science 257, 1142-1144, 1992
A:Title: Plakoglobin and beta-catenin: distinct but closely related.
A:Reference number: S35091; MUID:92376536
A:Accession: S35091
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-781 <BUT>
A:Cross-references: EMBL:M90364; NID:g192141; PIDN:AAA37280.1; PID:g192142
C:Keywords: cytoskeleton

Query Match 100.0%; Score 194; DB 2; Length 781;
Best Local Similarity 100.0%; Pred. No. 8.2e-18;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATKQEGMEGLGLTVQLVLLGSDDDINVTCAAGILSNLTC 39
|||||
Db 391 AATKQEGMEGLGLTVQLVLLGSDDDINVTCAAGILSNLTC 429

RESULT 4
T12689
armadillo segment polarity protein - fruit fly (Drosophila melanogaster)
N:Alternate names: protein 86E4.6
C:Species: Drosophila melanogaster
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 21-Jul-2000
C:Accession: T12689; A31861
R:Ferraz, C.; Vidal, S.; Brun, C.; Bucheton, A.; Demaille, J.G.
submitted to the EMBL Data Library, January 1998
A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.
A:Reference number: Z17572
A:Accession: T12689
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-843 <FER>
A:Cross-references: EMBL:AL021106; NID:e1371406; PID:e1249776; PIDN:CAAL5946.1
A:Experimental source: clone cosmid 63B12
R:Riggleman, B.; Wieschaus, E.; Schedl, P.
Genes Dev. 3, 96-113, 1989
A:Title: Molecular analysis of the Armadillo locus: uniformly distributed transcripts at
A:Reference number: A31861; MUID:89211895
A:Accession: A31861
A:Molecule type: DNA
A:Residues: 1-843 <RIC>
A:Cross-references: EMBL:X54468; NID:g7610; PIDN:CAA38350.1; PID:g7611
C:Genetics:

A:Gene: arm
A:Cross-references: FlyBase:FBgn0000117
A:Introns: 11/1; 557/3; 624/3; 715/3; 748/1
A:Note: 86E4.6
C:Keywords: cytoskeleton

Query Match 79.9%; Score 155; DB 2; Length 843;
Best Local Similarity 79.5%; Pred. No. 1.4e-12;
Matches 31; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 AATKQEGMEGLGLTVQLVLLGSDDDINVTCAAGILSNLTC 39
|||||
Db 399 AATKQEGMEGLGLTVQLVLLGSDDDINVTCAAGILSNLTC 437

RESULT 5
JC4835
beta-catenin - Hydra magnipapillata
C:Species: Hydra magnipapillata
C:Date: 15-Aug-1996 #sequence_revision 15-Oct-1996 #text_change 21-Jul-2000
C:Accession: JC4835
R:Hobmayer, E.; Hatta, M.; Fischer, R.; Fujisawa, T.; Holstein, T.W.; Sugiyama, T.
Gene 172, 155-159, 1996
A:Title: Identification of a Hydra homologue of the beta-catenin/plakoglobin/armadillo
A:Reference number: JC4835; MUID:96257271
A:Accession: JC4835
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-806 <HOB>
A:Cross-references: GB:U36781; NID:g1407600; PIDN:AAC47137.1; PID:g1407601
C:Comment: This protein plays roles in cadherin-mediated cell adhesion and in signal
C:Genetics:
A:Gene: betaCtn

Query Match 77.3%; Score 150; DB 2; Length 806;
Best Local Similarity 76.3%; Pred. No. 6.3e-12;
Matches 29; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 2 AATKQEGMEGLGLTVQLVLLGSDDDINVTCAAGILSNLTC 39
|||||
Db 451 AATKQEGMEGLGLTVQLVLLGSDDDINVTCAAGILSNLTC 488

RESULT 6
S33793
hypothetical protein - spoonworm (Urechis caupo)
C:Species: Urechis caupo
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Aug-1997
C:Accession: S33793
R:Rosenthal, E.
Biochim. Biophys. Acta 1173, 337-341, 1993
A:Title: Identification of homologues to beta-catenin/plakoglobin/armadillo in two in
A:Reference number: S33793; MUID:93305730
A:Accession: S33793
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-817 <ROS>

Query Match 75.8%; Score 147; DB 2; Length 817;
Best Local Similarity 76.9%; Pred. No. 1.6e-11;
Matches 30; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 AATKQEGMEGLGLTVQLVLLGSDDDINVTCAAGILSNLTC 39
|||||
Db 413 AATKSSDIEGLQLVLLGSDDDINVTCAAGILSNLTC 451

RESULT 7
S35093
plakoglobin - African clawed frog

C:Species: Xenopus laevis (African clawed frog)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jul-2000
C:Accession: S35093; S24636
R:Fouquet, B.; Zimbelmann, R.; Franke, W.W.
Differentiation 51, 187-194, 1992
A:Title: Identification of plakoglobin in oocytes and early embryos of Xenopus laevis: m
A:Reference number: S35093; MUID:93093332
A:Accession: S35093
A:Molecule type: mRNA
A:Residues: 1-738 <FOU>
A:Cross-references: EMBL:M95593; NID:9214656; PIDN:AAA49931.1; PID:g214657
R:Demarais, A.A.
submitted to the EMBL Data Library, June 1992
A:Reference number: S24636
A:Accession: S24636
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 133-184, 'V', 186-225, 'T', 227-292 <DEM>
A:Cross-references: EMBL:X67078; NID:965252; PID:g65253
C:Keywords: cytoskeleton

Query Match 67.5%; Score 131; DB 2; Length 738;
Best Local Similarity 65.8%; Pred. No. 2e-09;
Matches 25; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 ATKQSGMGLLGLTVQLGSDDDINVTCAAGILSLNLT 39
|||||: : || || |||||: |||||
DB 379 ATKQGLDNVLKILVQLNSDDVNVLTATGTLNLT 416

RESULT 8

A32905
Plakoglobin, desmosomal - human
C:Species: Homo sapiens (man)
C:Date: 22-Nov-1989 #sequence_revision 22-Nov-1989 #text_change 29-Aug-1997
C:Accession: A32905
R:Frankel, W.W.; Goldschmidt, M.D.; Zimbelmann, R.; Mueller, H.M.; Schiller, D.L.; Cowin,
Proc. Natl. Acad. Sci. U.S.A. 86, 4027-4031, 1989
A:Title: Molecular cloning and amino acid sequence of human plakoglobin, the common junct
A:Reference number: A32905; MUID:89264555
A:Accession: A32905
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-744 <ERA>
A:Cross-references: GB:M23410
C:Genetics:
A:Gene: GDB:JUP
A:Cross-references: GDB:L26565; OMIM:173325
A:Map position: 7pter-7qter
C:Keywords: cytoskeleton

Query Match 66.0%; Score 128; DB 2; Length 744;
Best Local Similarity 65.8%; Pred. No. 5e-09;
Matches 25; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 2 ATKQSGMGLLGLTVQLGSDDDINVTCAAGILSLNLT 39
|||||: : || || |||||: |||||
DB 382 ATKQGLSVLKILVQLNSDDVNVLTATGTLNLT 419

RESULT 9

S33794
hypothetical protein - sea urchin (Tripneustes gratilla)
C:Species: Tripneustes gratilla
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Aug-1997
C:Accession: S33794
R:Rosenthal, E.
Biochim. Biophys. Acta 1173, 337-341, 1993
A:Title: Identification of homologues to beta-catenin/plakoglobin/armadillo in two inver
A:Reference number: S33793; MUID:93305730

A:Accession: S33794
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-820 <ROS>

Query Match 61.9%; Score 120; DB 2; Length 820;
Best Local Similarity 57.3%; Pred. No. 6.5e-08;
Matches 25; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 6 ECGMGLLGLTVQLGSDDDINVTCAAGILSLNLT 39
|||||: : || || |||||: |||||
DB 412 DDIEPLQLVQLASNDINVTCAAGILSLNLT 445

RESULT 10

S35092
plakoglobin - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 17-Mar-1999
C:Accession: S35092
R:Butz, S.; Stappert, J.; Weissig, H.; Kemler, R.
Science 257, 1142-1144, 1992
A:Title: Plakoglobin and beta-catenin: distinct but closely related.
A:Reference number: S35091; MUID:92376536
A:Accession: S35092
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-621 <BUT>
A:Cross-references: EMBL:M90365
C:Keywords: cytoskeleton

Query Match 58.2%; Score 113; DB 2; Length 621;
Best Local Similarity 62.2%; Pred. No. 4.2e-07;
Matches 23; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 2 ATKQSGMGLLGLTVQLGSDDDINVTCAAGILSLNLT 38
|||||: : || || |||||: |||||
DB 259 ATKQGLSVLKILVQLNSDDVNVLTATGTLNLT 295

RESULT 11

T23341
beta-catenin - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
C:Accession: T23341; T42221
R:Harris, B.
submitted to the EMBL Data Library, November 1996
A:Reference number: 219729
A:Accession: T23341
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-678 <WIL>
A:Cross-references: EMBL:Z81564; PIDN:CAB04572.1; GSPDB:GN00019; CESP:K05C4.6
A:Experimental source: clone K05C4
R:Costa, M.; Raich, W.; Agbunag, C.; Hardin, J.; Pries, J.R.
submitted to the EMBL Data Library, July 1997
A:Description: A putative catenin-cadherin system mediates morphogenesis of the C. el
A:Reference number: 222085
A:Accession: T42221
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-678 <COS>
A:Cross-references: EMBL:AF016853; PIDN:AAB94552.1
C:Genetics:
A:Gene: hmp-2; CESP:K05C4.6
A:Map position: 1
A:Introns: 14/1; 235/1; 290/1; 620/3

Query Match 36.6%; Score 71; DB 2; Length 678;

```

Best Local Similarity 35.9%; Pred. NO.0.19;
Matches 14; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

QY 1 AATKQEGMGLLTGLVQLLGSDDDINVTCAAGILSLNLT 39
      ||| : | : | : : : ||| |||||
Db 321 SATNEENLTQLIKLEIRVANGDMTACAGTSLNLT 359

RESULT 12
T03746
hypothetical protein wrm-1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 21-Jul-2000
C:Accession: T03746
R:R;Rocheleau, C.E.; 1997
Cell 90, 707-716, 1997
A:title: Wnt signaling and an APC related gene specify endoderm in early C. elegans embryo
A:Reference number: Z15051; MUID:97433081
A:Accession: T03746
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-796 <R0C>
A:Cross-references: EMBL:AF013951; NID:g2343264; PIDN:AAC47748.1; PID:g2343265
A:Experimental source: strain Bristol
C:Genetics:
A:Gene: wrm-1
A:Map position: 3

```

```

Query Match          33.0%; Score 64; DB 2; Length 796;
Best Local Similarity 33.3%; Pred. No. 1.9;
Matches 12; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY      2 ATKQEGEGLLGTIVQLLGSDDINVTCAAGILSNL 37
      | : : : | : : : | : : | : | : | :
Db      454 AIQEQDKMSIHCIVQLIGCSDDTIVELATGLTRNI 489

RESULT 13
TI5342
hypothetical protein B0336.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: TI5342
R:Taich, A.
submitted to the EMBL Data Library, July 1995
A:Description: The sequence of C. elegans cosmid B0336.
A:Reference number: Z18331
A:Accession: TI5342
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-964 <TAI>
A:Cross-references: EMBL:U32305; NID:g912752; PID:g912753; PIDN:AAC46827.1; CESP:B0336.
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:B0336.1
A:Introns: 132/3; 152/1; 188/1; 245/3; 373/3; 667/1; 782/3; 838/2; 906/1

```

```
Query Match      33.0%; Score 64; DB 2; Length 964;  
Best Local Similarity 33.3%; Pred. NO. 2.3;  
Matches 12; Conservative 10; Mismatches 14; Indels 0; Gaps 0;
```

OY 2 ATKQEGMGLLTFLVQLLGSDDINVVTCAGILSNL 37
| : : | : : : ||||| : : : | | | :
DB 597 ATIQEQDKESIHCIVOLIGCSDVITVELATGTLRNI 632

RESULT 14
A86366
T76J12.6 protein - Arabidopsis thaliana
C Species: Arabidopsis thaliana (mouse-ear cress)
C Date: 02-Mar-2001 #sequence revision 02-Mar-2001 #text change 27-Nov-2001

C:Accession: A86366
C:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luos, J.S.; Maiti, R.; Marzilia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: A86366
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1299 <STO>
A:Cross-references: GB:AE005172; NID:g2829894; PIDN:AAC00602.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

	Query Match	Score	DB 2;	Length
Best Local Similarity	44.4%	Pred. No. 5.7;	Indels 0;	Gaps 0;
Matches	12: Conservative	5: Mismatches	10: Indels	0: Gaps

RESULT **15**

AEL961.
Zam protein [imported] - Anabaena sp. (strain PCC 7120)
C:Species: Anabaena sp.
A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
C:Accession: AEL961
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigoien, J.M.; Shimizu, K.; Taniguchi, M.; Tabatake, H.; Takazawa, M.; Yamada, M.; Sakagami, Y.; Sugimoto, M.; Nakazaki, I.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabatani, R.E.; Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigoien, J.M.; Shimizu, K.; Taniguchi, M.; Tabatake, H.; Takazawa, M.; Yamada, M.; Sakagami, Y.; Sugimoto, M.; Nakazaki, I.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabatani, R.E.

DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena azotofixans
A:Reference number: AB1807; PMID:11759840
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-783 <KUR>
A:Cross-references: GB:BAO00019; PIDN:BAB73197.1; PID:g17130587; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: zam

```

Query Match      29.1%; Score 56.5; DB 2; Length 783;
Best Local Similarity 43.3%; Restrict No. 19;
Matches: 13; Conservative    6; Mismatches   6; Indels   5; Gaps   1;

QY  12 LGTLVOLLASD-----DINVVTCAAGILSN 36
     .:::|::::|::||::|::|::|::|::|::|::|
ph  192 VGRVVVLGSDAEAAADIDLVTCCKHDLN 221

```

Search completed: July 29, 2002, 16:09:16
Job time: 726 sec

THIS PAGE BLANK (USPST)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 16:09:54 ; Search time 34.24 Seconds
(without alignments)
44.102 Million cell updates/sec

Title: US-09-641-104A-10

Perfect score: 194

Sequence: 1 AATKQEGMEGLLTGLVQLLGSDDDINVTCAAGILSNLTC 39

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	194	100.0	781	1 CTNB_HUMAN	P35222 homo sapien
2	194	100.0	781	1 CTNB_MOUSE	Q02248 mus musculus
3	194	100.0	781	1 CTNB_RAT	Q9wu82 rattus norv
4	194	100.0	781	1 CTNB_XENLA	P26233 xenopus lae
5	155	79.9	843	1 ARM_DROME	P18824 drosophila
6	147	75.8	818	1 CTNB_URECA	P35224 urechis cau
7	144	74.2	813	1 ARM_MUSDO	Q02453 musca domes
8	131	67.5	738	1 PLAK_XENLA	P30998 xenopus lae
9	128	66.0	621	1 PLAK_MOUSE	Q02257 mus musculus
10	128	66.0	743	1 PLAK_HUMAN	P14923 homo sapien
11	120	61.9	820	1 CTNB_TRIGR	P35223 tripeustes
12	56	28.9	322	1 Y787_SYNY3	Q55946 synechocyst
13	54.5	28.1	868	1 AC02_SYNY3	P74582 synechocyst
14	54	27.8	525	1 SERA_BACSU	P35136 bacillus su
15	53	27.3	554	1 SERA_SYNY3	P73821 synechocyst
16	52.5	27.1	1008	1 MMLA_MYCLE	Q49619 mycobacteri
17	52	26.8	238	1 Y933_METJA	Q58343 methanococ
18	51.5	26.5	487	1 VPL_HAEIN	P44233 haemophilus
19	51.5	26.5	578	1 VATA_METBA	P22662 methanosarc
20	50.5	26.0	80	1 GAEL_RANRU	P80398 rana rugosa
21	50	25.8	529	1 IMA2_HUMAN	P52292 homo sapien
22	50	25.8	529	1 IMA2_MOUSE	P52293 mus musculus
23	50	25.8	727	1 PKP1_BOVIN	Q28161 bos taurus
24	50	25.8	728	1 PKP1_MOUSE	P97350 mus musculus
25	50	25.8	747	1 PKP1_HUMAN	Q13835 homo sapien
26	49.5	25.5	313	1 FAJF_HAEIN	P44554 haemophilus
27	49.5	25.5	473	1 AK_METJA	Q57991 methanococ
28	49	25.3	108	1 END4_THENE	O86954 thermotoga
29	49	25.3	215	1 Y267_BUCAI	P57355 buchnera ap
30	49	25.3	287	1 END4_THEMEA	Q9wyj7 thermotoga
31	49	25.3	292	1 DHYS_AERPE	Q9ye77 aeropyrum p
32	49	25.3	297	1 HTPX_STRGC	O30795 streptococ
33	49	25.3	483	1 PANF_ECOLI	P16256 escherichia

34	49	25.3	514	1 IMA_CAEEL	Q19969 caenorhabdi
35	49	25.3	875	1 IMB1_RAT	P52296 rattus norv
36	49	25.3	876	1 IMB1_HUMAN	Q14974 homo sapien
37	49	25.3	876	1 IMB1_MOUSE	P70168 mus musculus
38	48.5	25.0	1012	1 POLS_IBDVC	P15480 avian infec
39	48	24.7	276	1 YXKO_BACSU	P94368 bacillus su
40	48	24.7	527	1 PPOX_PROFR	O32434 propionibac
41	48	24.7	558	1 GDSL_BOVIN	Q04173 bos taurus
42	48	24.7	558	1 GDSL_HUMAN	P52306 homo sapien
43	48	24.7	702	1 SPFL_ARATH	O9si64 arabidopsis
44	47.5	24.5	491	1 NUOM_RICPR	Q9sc90 rickettsia
45	47.5	24.5	493	1 NUOM_RICCN	Q94996 rickettsia

ALIGNMENTS

RESULT 1

ID	CTNB_HUMAN	STANDARD	PRT	781 AA
AC	P35222			
DT	01-FEB-1994 (Rel. 28, Created)			
DT	01-FEB-1994 (Rel. 28, Last sequence update)			
DE	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Beta-catenin.			
GN	CTNNB1 OR CTNNB.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Placenta;			
RX	MEDLINE=95105247; PubMed=7806582;			
RA	Huelsen J., Birnmeier W., Behrens J.;			
RT	"E-cadherin and APC compete for the interaction with beta-catenin and the cytoskeleton.";			
RL	J. Cell Biol. 127:2061-2069(1994).			
RN	[2]			
RP	REVIEW			
RA	MEDLINE=20145417; PubMed=10679188;			
RX	Kikuchi A.;			
RT	"Regulation of beta-catenin signaling in the Wnt pathway.";			
RL	Biochem. Biophys. Res. Commun. 268:243-248(2000).			
RN	[3]			
RP	X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 133-664.			
RX	MEDLINE=20578806; PubMed=11136974;			
RA	Graham T.A., Weaver C., Mao F., Kimmelman D., Xu W.;			
RL	"Crystal structure of a beta-catenin/Tcf complex.";			
Cell	103:885-896(2000).			
CC	-1- FUNCTION: INVOLVED IN THE REGULATION OF CELL ADHESION AND IN SIGNAL TRANSDUCTION THROUGH THE WNT PATHWAY.			
CC	-1- SUBUNIT: THE CYTOSOLIC FORM BOUND TO AXIN ALLOWS THE STIMULATION OF ITS DEGRADATION. THE CYTOSOLIC FORM ALSO INTERACTS WITH APC AND E-CADHERIN. THE NUCLEAR FORM BINDS TCF/LEF-1 AND MAY ALSO BIND PONTIN52 AND DUPLIN. CAN HETERODIMERIZE WITH ALPHA-CATENIN AND ALSO FORM A TRIMERIC COMPLEX WITH CADHERIN. MEDIATES THE ANCHORAGE OF THE CADHERIN COMPLEX TO THE CORTICAL ACTIN CYTOSKELETON.			
CC	-1- SUBCELLULAR LOCATION: CYTOPLASMIC WHEN IT IS UNSTABILIZED (HIGH LEVEL OF PHOSPHORYLATION). TRANSLOCATES TO THE NUCLEUS WHEN IT IS STABILIZED (LOW LEVEL OF PHOSPHORYLATION).			
CC	-1- PTM: PROBABLY PHOSPHORYLATED BY GSK-3B. WNT-DEPENDENT ACTIVATION OF DVL ANTAGONIZES THE ACTION OF GSK-3B. PHOSPHORYLATION OF BETA-CATENIN IS REDUCED AND THE PROTEIN IS NO LONGER DEGRADED, RESULTING IN ITS ACCUMULATION IN CYTOPLASM.			
CC	-1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.			
CC	-1- SIMILARITY: CONTAINS 12 ARM REPEATS.			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC EMBL; X87838: CAA61107.1; -.
CC EMBL; Z19054: CAA79497.1; -.
CC DR PIR; S31988: S31988.
CC DR PDB; 1G3J; 17-JAN-01.
CC DR TRANSFAC; T02872; -.
CC DR MIM; 116806; -.
CC DR InterPro; IPR000225; Armadillo.
CC DR Pfam; PF00514; Armadillo_seg; 12.
CC DR SMART; SM00185; ARM; 11.
CC DR PROSITE; PS50176; ARM_REPEAT; 9.
CC KW Cell adhesion; Cytoskeleton; Structural protein; Nuclear protein;
KW Repeat; 3D-structure.
FT REPEAT 151 191 ARM 1.
FT REPEAT 193 234 ARM 2.
FT REPEAT 235 276 ARM 3.
FT REPEAT 277 318 ARM 4.
FT REPEAT 319 360 ARM 5.
FT REPEAT 361 389 ARM 6.
FT REPEAT 400 441 ARM 7.
FT REPEAT 442 484 ARM 8.
FT REPEAT 489 530 ARM 9.
FT REPEAT 531 571 ARM 10.
FT REPEAT 594 636 ARM 11.
FT REPEAT 637 666 ARM 12.
SQ SEQUENCE 781 AA; 85496 MW; C878F165A3EEF86E CRC64;

Query Match 100.0%; Score 194; DB 1; Length 781;
Best Local Similarity 100.0%; Pred. No. 5.7e-18;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATKQEGMGLLGLTVOLLGSDDDINVTCAAGILSNLTC 39
    |||||
DB 391 AATKQEGMGLLGLTVOLLGSDDDINVTCAAGILSNLTC 429

RESULT 2
CTNB_MOUSE STANDARD; PRT; 781 AA.
AC Q02248;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Beta-catenin.
GN CTNNB1 OR CATNB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92376536; PubMed=1509266;
RA Butz S., Stappert J., Weissig H., Kemler R.;
RT "Plakoglobin and beta-catenin: distinct but closely related.";
RL Science 257:1142-1144(1992).
RN [2]
RP REVIEW.
RX MEDLINE=20145417; PubMed=10679188;
RA Kikuchi A.;
RT "Regulation of beta-catenin signaling in the Wnt pathway.";
RL Biochem. Biophys. Res. Commun. 268:243-248(2000).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 150-665.
RX MEDLINE=97442350; PubMed=9298899;
RA Huber A.H., Nelson W.J., Weiss W.I.;
RT "Three-dimensional structure of the armadillo repeat region of beta-
    catenin.";
RL Cell 90:871-882(1997).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 118-149 IN COMPLEX WITH

```

```

CTNNB1.
MEDLINE=20337986; PubMed=10882138;
RX Pokutta S., Weis W.I.;
RA "Structure of the dimerization and beta-catenin-binding region of
    alpha-catenin.";
RL Mol. Cell 5:533-543(2000).
CC -1- FUNCTION: INVOLVED IN THE REGULATION OF CELL ADHESION AND IN
    SIGNAL TRANSDUCTION THROUGH THE WNT PATHWAY.
CC -1- SUBUNIT: THE CYTOSOLIC FORM BOUND TO AXIN ALLOWS THE STIMULATION
    OF ITS DEGRADATION. THE CYTOSOLIC FORM ALSO INTERACTS WITH APC AND
    E-CADHERIN. THE NUCLEAR FORM BINDS TCF/LEF-1 AND MAY ALSO BIND
    PONTIN52 AND DUPLIN. CAN HETERODIMERIZE WITH ALPHA-CATENIN AND
    ALSO FORM A TRIMERIC COMPLEX WITH CADHERIN. MEDIATES THE ANCHORAGE
    OF THE CADHERIN COMPLEX TO THE CORTICAL ACTIN CYTOSKELETON.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC WHEN IT IS UNSTABILIZED (HIGH
    LEVEL OF PHOSPHORYLATION). TRANSLOCATES TO THE NUCLEUS WHEN IT IS
    STABILIZED (LOW LEVEL OF PHOSPHORYLATION).
CC -1- PTM: PROBABLY PHOSPHORYLATED BY GSK-3B. WNT-DEPENDENT ACTIVATION
    OF DVL ANTAGONIZES THE ACTION OF GSK-3B. PHOSPHORYLATION OF BETA-
    CATENIN IS REDUCED AND THE PROTEIN IS NO LONGER DEGRADED,
    RESULTING IN ITS ACCUMULATION IN CYTOPLASM.
CC -1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -1- SIMILARITY: CONTAINS 12 ARM REPEATS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
    between the Swiss Institute of Bioinformatics and the EMBL outstation -
    the European Bioinformatics Institute. There are no restrictions on its
    use by non-profit institutions as long as its content is in no way
    modified and this statement is not removed. Usage by and for commercial
    entities requires a license agreement (See http://www.isb-sib.ch/announce/
    or send an email to license@isb-sib.ch).
```

```

EMBL; M90364; AAA37280.1; -.
PIR; S35091; S35091.
PDB; 2BCT; 15-OCT-97.
PDB; 3BCT; 19-NOV-97.
PDB; IDOW; 12-JUL-00.
DR TRANSFAC; T02984; -.
DR MGI; MGI:88276; Catnb.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 12.
DR SMART; SM00185; ARM; 11.
DR PROSITE; PS50176; ARM_REPEAT; 9.
DR Cell adhesion; Cytoskeleton; Structural protein; Nuclear protein;
KW Repeat; 3D-structure.
FT REPEAT 151 191 ARM 1.
FT REPEAT 193 234 ARM 2.
FT REPEAT 235 276 ARM 3.
FT REPEAT 277 318 ARM 4.
FT REPEAT 319 360 ARM 5.
FT REPEAT 361 389 ARM 6.
FT REPEAT 400 441 ARM 7.
FT REPEAT 442 484 ARM 8.
FT REPEAT 489 530 ARM 9.
FT REPEAT 531 571 ARM 10.
FT REPEAT 594 636 ARM 11.
FT REPEAT 637 666 ARM 12.
SQ SEQUENCE 781 AA; 85470 MW; D708F170A3FBED6E CRC64;

Query Match 100.0%; Score 194; DB 1; Length 781;
Best Local Similarity 100.0%; Pred. No. 5.7e-18;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATKQEGMGLLGLTVOLLGSDDDINVTCAAGILSNLTC 39
    |||||
DB 351 AATKQEGMGLLGLTVOLLGSDDDINVTCAAGILSNLTC 429

RESULT 3
CTNB_RAT STANDARD; PRT; 781 AA.
ID CTNB_RAT
AC Q9WU22;

```

16-OCT-2001 (Rel. 40, Created)
 16-OCT-2001 (Rel. 40, Last sequence update)
 01-MAR-2002 (Rel. 41, Last annotation update)
 Beta-catenin.
 GN CTNNB1.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE-Testis;
 RX MEDLINE-95428593; PubMed=10497305;
 RA Chung S.S.W., Lee W.M., Cheng C.Y.;
 RT "Study on the formation of specialized inter-Sertoli cell junctions in
 RT vitro.";
 RL J. Cell. Physiol. 181:258-272(1999).
 CC -!- FUNCTION: INVOLVED IN THE REGULATION OF CELL ADHESION AND IN
 CC -!- SIGNAL TRANSDUCTION THROUGH THE WNT PATHWAY (BY SIMILARITY).
 CC -!- SUBUNIT: THE CYTOSOLIC FORM BOUND TO AXIN ALLOWS THE STIMULATION
 CC OF ITS DEGRADATION. THE CYTOSOLIC FORM ALSO INTERACTS WITH APC AND
 CC E-CADHERIN. THE NUCLEAR FORM BINDS TCF/LEF-1 AND MAY ALSO BIND
 CC PONTIN52 AND DUPLIN. CAN HETERODIMERIZE WITH ALPHA-CATENIN AND
 CC ALSO FORM A TRIMERIC COMPLEX WITH CADHERIN. MEDIATES THE ANCHORAGE
 CC OF THE CADHERIN COMPLEX TO THE CORTICAL ACTIN CYTOSKELETON (BY
 CC SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC WHEN IT IS UNSTABILIZED (HIGH
 CC LEVEL OF PHOSPHORYLATION). TRANSLOCATES TO THE NUCLEUS WHEN IT IS
 CC STABILIZED (LOW LEVEL OF PHOSPHORYLATION) (BY SIMILARITY).
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE TESTIS.
 CC -!- DEVELOPMENTAL STAGE: HIGHLY EXPRESSED AT E30-60 DAY DPC IN THE
 CC TESTIS. REDUCED EXPRESSION AT E90 DAY DPC.
 CC -!- PTM: PROBABLY PHOSPHORYLATED BY GSK-3B. WNT-DEPENDENT ACTIVATION
 CC OF DVL ANTAGONIZES THE ACTION OF GSK-3B. PHOSPHORYLATION OF BETA-
 CC CATENIN IS REDUCED AND THE PROTEIN IS NO LONGER DEGRADED.
 CC RESULTING IN ITS ACCUMULATION IN CYTOPLASM (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
 CC -!- SIMILARITY: CONTAINS 12 ARM REPEATS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF121265; AAD28504.1; --
 CC HSPF: Q02248; 2BCT.
 CC InterPro: IPR000225; Armadillo.
 CC Pfam: PF00514; Armadillo_seg; 12.
 CC SMART: SM00185; ARM; 11.
 CC PROSITE: PS50176; ARM_REPEAT; 9.
 CC Cell adhesion; Cytoskeleton; Structural protein; Nuclear protein;
 KW Repeat.
 FT REPEAT 151 191 ARM 1.
 FT REPEAT 193 234 ARM 2.
 FT REPEAT 235 276 ARM 3.
 FT REPEAT 277 318 ARM 4.
 FT REPEAT 319 360 ARM 5.
 FT REPEAT 361 389 ARM 6.
 FT REPEAT 400 441 ARM 7.
 FT REPEAT 442 484 ARM 8.
 FT REPEAT 489 530 ARM 9.
 FT REPEAT 531 571 ARM 10.
 FT REPEAT 594 636 ARM 11.
 FT REPEAT 637 666 ARM 12.
 SQ SEQUENCE 781 AA; 85454 MW; 9C29186BD54B87 CRC64;

Query Match 100.0%; Score 194; DB 1; Length 781;
 Best Local Similarity 100.0%; Pred. No. 5.7e-18;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATKQEGMEGLLGLTVQLLGSDDINVTCAAGILSNLTC 39
 DB 391 AATKQEGMEGLLGLTVQLLGSDDINVTCAAGILSNLTC 429
 RESULT 4
 CTNNB_XENLA6 STANDARD; PRT; 781 AA.
 ID AC P26233;
 DT 01-MAY-1992 (Rel. 22; Created)
 DT 01-MAY-1992 (Rel. 22; Last sequence update)
 DT 01-MAR-2002 (Rel. 41; Last annotation update)
 DE Beta-catenin.
 OS Xenopus laevis; (African clawed frog).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
 CC Xenopodidae; Xenopus.
 CC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE-Brain;
 RX MEDLINE-92073903; PubMed-1962194;
 RA McCrea P.D., Turk C.W., Gumbiner B.M.;
 RT "A homolog of the armadillo protein in Drosophila (plakoglobin)
 RT associated with E-cadherin.";
 RL Science 254:1359-1361(1991).
 CC -!- FUNCTION: BINDS TO THE CYTOPLASMIC DOMAIN OF THE CELL-CELL
 CC ADHESION MOLECULE E-CADHERIN, AND PERHAPS TO OTHER (MEMBRANE)
 CC PROTEINS. THE ASSOCIATION OF CATENINS TO CADHERINS PRODUCES A
 CC COMPLEX WHICH IS LINKED TO THE ACTIN FILAMENT NETWORK, AND WHICH
 CC SEEMS TO BE OF PRIMARY IMPORTANCE FOR CADHERINS CELL-ADHESION
 CC PROPERTIES. ESSENTIAL COMPONENT OF THE VERTEBRATE WNT-SIGNALING
 CC PATHWAY AND IS REQUIRED FOR FORMATION OF THE EMBRYONIC AXIS.
 CC -!- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
 CC -!- SIMILARITY: CONTAINS 10 ARM REPEATS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M70113; AAA49670.1; --
 CC PIR: S35099; S35099.
 CC HSPF: Q02248; 2BCT.
 CC TRANSFAC: T03026; --
 CC InterPro: IPR000225; Armadillo.
 CC Pfam: PF00514; Armadillo_seg; 12.
 CC SMART: SM00185; ARM; 11.
 CC PROSITE: PS50176; ARM_REPEAT; 9.
 CC Cell adhesion; Cytoskeleton; Structural protein; Repeat.
 KW Repeat.
 FT REPEAT 141 180 ARM 1.
 FT REPEAT 181 225 ARM 2.
 FT REPEAT 226 264 ARM 3.
 FT REPEAT 267 306 ARM 4.
 FT REPEAT 307 350 ARM 5.
 FT REPEAT 351 390 ARM 6.
 FT REPEAT 391 429 ARM 7.
 FT REPEAT 430 473 ARM 8.
 FT REPEAT 474 519 ARM 9.
 FT REPEAT 520 562 ARM 10.
 FT REPEAT 563 603 ARM 11.
 FT REPEAT 604 644 ARM 12.
 SQ SEQUENCE 781 AA; 85449 MW; 3ECD2723239F799 CRC64;

Query Match 100.0%; Score 194; DB 1; Length 781;
 Best Local Similarity 100.0%; Pred. No. 5.7e-18;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATKQEGMEGLLGLTVQLLGSDDINVTCAAGILSNLTC 39
 DB 391 AATKQEGMEGLLGLTVQLLGSDDINVTCAAGILSNLTC 429

DB 391 AATKQEGMEGLCTVQLGSDINNVTCAGILSNLTC 429

RESULT 5

ARM_DROME STANDARD; PRT; 843 AA.

ID ARM_DROME

AC P18824: 002371; 09W546;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DE 01-MAR-2002 (Rel. 41, Last annotation update)

DE Armadillo segment polarity protein.

ARM OR EG:86E4.6 OR CG11579.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

[1]

RP SEQUENCE FROM N.A. (CYTOPLASMIC ISOFORM).

RC STRAIN=OREGON-R;

RX MEDLINE=89211895; PubMed=2707602;

RA Riggelman B., Wieschaus E., Schedl P.;

RT "Molecular analysis of the Armadillo locus: uniformly distributed transcripts and a protein with novel internal repeats are associated with a Drosophila segment polarity gene.";

RT Genes Dev. 3:96-113(1989).

RL [2]

RN SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, AND ALTERNATIVE SPLICING.

RP TISSUE=Head;

RC MEDLINE=98298928; PubMed=9635189;

RX Loureiro J., Peifer M.;

RA "Roles of Armadillo, a Drosophila catenin, during central nervous system development.";

RT Curr. Biol. 8:622-632(1998).

RL [3]

RN SEQUENCE FROM N.A. (CYTOPLASMIC ISOFORM).

RP STRAIN=BERKELEY;

RC MEDLINE=20196011; PubMed=10731137;

RX Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D., Anantides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers J.H., Blazej R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D., Ballew R.M., Basu A., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berhan B.P., Bhandari D., Bolshakov S., Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P., Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W., Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K., Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C., Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacileb J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spler E., Spradling A.C., Stapleton M., Strong R., Sun E., Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

"The genome sequence of Drosophila melanogaster.";

Science 287:2185-2195(2000).

[4]

RP SEQUENCE FROM N.A. (CYTOPLASMIC ISOFORM).

RC STRAIN=OREGON-R;

RX MEDLINE=20196011; PubMed=10731137;

RA Benos P.V., Gatt M.K., Ashburner M., Murphy L., Harris D., Barreil B.G., Ferraz C., Vidal S., Brun C., Demallies J., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S., Galibert F., Borkova D., Minana B., Kafatos F.C., Louis C., Siden-Kiamos I., Bolshakov S., Papagiannakis G., Spanos L., Cox S., Madueno E., de Pablos B., Modolell J., Peter A., Schoettler P., Werner M., Mourikoti F., Beilert N., Dowe G., Schaefer U., Jaekle H., Bucheton A., Callister D.M., Campbell L.A., Darlamitsou A., Henderson N.S., McMillan P.J., Salles C., Tait E.A., Valenti P., Saunders R.D.C., Glover D.M.;

RT "From sequence to chromosome: the tip of the X chromosome of D. melanogaster.";

RT Science 287:2220-2222(2000).

RL [5]

RN PHOSPHORYLATION.

RP MEDLINE=95113174; PubMed=7529201;

RX Peifer M., Pai L.-M., Casey M.;

RT "Phosphorylation of the Drosophila adherens junction protein Armadillo: roles for wingless signal and zeste-white 3 kinase.";

RT Dev. Biol. 166:543-556(1994).

CC -!- FUNCTION: NEURAL ISOFORM MAY ASSOCIATE WITH CADN AND PARTICIPATE IN THE TRANSMISSION OF DEVELOPMENTAL INFORMATION. CAN ASSOCIATE WITH ALPHA-CATENIN. CYTOPLASMIC ISOFORM ACCUMULATES THROUGH WG SIGNALING. ARM FUNCTION IN WG SIGNAL TRANSDUCTION IS REQUIRED EARLY IN DEVELOPMENT FOR DETERMINATION OF NEUROBLAST FATE. ARM AND ABL PROTEINS FUNCTION COOPERATIVELY AT ADHERENS JUNCTIONS IN BOTH THE CNS AND EPIDERMIS.

CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC AND MEMBRANE-ASSOCIATED WITH THE INNER SURFACE OF CELL MEMBRANE.

CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; CYTOPLASMIC (SHOWN HERE) AND NEURAL; ARE PRODUCED BY ALTERNATIVE SPLICING.

CC -!- TISSUE SPECIFICITY: CYTOPLASMIC ISOFORM IS PREDOMINANT BEFORE GERM BAND RETRACTION. AFTER RETRACTION AND DURING LARVAL STAGES, IT IS FOUND IN HIGH LEVELS IN SPECIFIC CELLS ALONG THE CNS MIDLINE.

CC NEURAL ISOFORM IS FIRST SEEN AFTER GERM BAND RETRACTION IN THE AXON TRACTS OF THE CNS, ALSO PRESENT IN AXONS DURING LARVAL STAGES AND ACCUMULATES IN THE MOTOR NEURONS OF THE SEGMENTAL AND INTERSEGMENTAL NERVES AS THEY EXIT THE CNS. BOTH ISOFORMS ACCUMULATE IN THE PNS.

CC -!- DEVELOPMENTAL STAGE: PRESENT AT ALL STAGES, BUT REACHED THE HIGHEST LEVELS DURING EARLY TO MID-EMBRYOGENESIS.

CC -!- PTM: PHOSPHORYLATED ON SER, THR AND TYR RESIDUES. LEVEL OF PHOSPHORYLATION VARIES BOTH DURING EMBRYONIC DEVELOPMENT AND FROM EMBRYONIC TISSUE TO TISSUE. SGG IS REQUIRED FOR PHOSPHORYLATION AND WG SIGNAL NEGATIVELY REGULATES ARM PHOSPHORYLATION.

CC -!- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.

CC -!- SIMILARITY: CONTAINS 12.5 ARM REPEATS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; X54468; CAA38350.1; --

EMBL; AF001213; AAB58731.1; --

EMBL; AE003422; AAF45688.1; ALT_INIT.

EMBL; AL021106; CAA15946.1; --

EMBL; AL021086; CAA15946.1; JOINED.

EMBL; AL021086; CAA15935.1; --

EMBL; AL021106; CAA15935.1; JOINED.

PIR; A31861; A31861.

OC Perlygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Muscoidea; Muscidae; Musca.
NCBI_TaxID=7370;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=93247062; PubMed=8483160;
RA Peifer M.A., Wieschaus E.;
RT "The product of the Drosophila melanogaster segment polarity gene
RT armadillo is highly conserved in sequence and expression in the
RT housefly Musca domestica";
RL J. Mol. Evol. 36:224-233(1993).
CC -I- FUNCTION: SEGMENT POLARITY PROTEIN. MUTATION IN ARM PRIMARILY
CC AFFECT THE POSTERIOR PART OF THE SEGMENT AND LEAD TO THE
CC PRODUCTION OF ANTERIOR STRUCTURES WITHIN THIS REGION. MAY BIND
CC A CADHERIN AND PARTICIPATE IN THE TRANSMISSION OF DEVELOPMENTAL
CC INFORMATION.
CC -I- SUBCELLULAR LOCATION: INNER SURFACE OF CELL MEMBRANE.
CC -I- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -I- SIMILARITY: CONTAINS 10 ARM REPEATS.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use: by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

DR EMBL: L04874; AAA29292.1; -
DR HSPG: Q02248; IDOW
DR InterPro: IPR000225; Armadillo.
DR Pfam: PF00514; Armadillo_seg; 12.
DR SMART: SM00185; ARM; 11.

entitles requires a license agreement (See <http://www.isb-sib.ch/announcement> or send an email to license@isb-sib.ch).

CC EMBL: M95593; AAA49931.1; -
CC EMBL: X67078; CAA47463.1; -
CC PIR: S24636; S24636.
CC PIR: S35093; S35093.
CC HSSP: Q02248; 2BCT.
CC InterPro: IPR000225; Armadillo.
CC Pfam: PF00514; Armadillo_seg; 11.
CC SMART: SM00185; ARM; 10.
CC PROSITE: PS0176; ARM_REPEAT; 9.
CC Cell Adhesion: Cytoskeleton; Structural protein; Repeat.
FT REPEAT 128 167 ARM 1.
FT REPEAT 212 251 ARM 2.
FT REPEAT 254 293 ARM 3.
FT REPEAT 338 377 ARM 4.
FT REPEAT 379 416 ARM 5.
FT REPEAT 419 460 ARM 6.
FT REPEAT 466 506 ARM 7.
FT REPEAT 508 547 ARM 8.
FT REPEAT 570 609 ARM 9.
FT REPEAT 611 650 ARM 10.
FT CONFLICT 185 185 I -> V (IN REF. 2).
FT CONFLICT 226 226 A -> T (IN REF. 2).
SQ SEQUENCE 738 AA; 81711 MW; 569DBE69D08BBC58 CRC64;

Query Match 67.5%; Score 131; DB 1; Length 738;
Best Local Similarity 65.8%; Pred. No. 1.e-09;
Matches 25; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 ATKQEGMGLLTGLVQLGSDDDINVTCAAGILSLNLT 39
DB 379 ATKQGLDNLVQLGSDDDINVTCAAGILSLNLT 416
RESULT 9
PLAK_MOUSE
ID PLAK_MOUSE STANDARD; PRT; 621 AA.
AC Q02257;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Junction plakoglobin (Desmoplakin III) (Fragment).
GN JUP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93376536; Pubmed=1509266;
RA Butz S., Stappert J., Weissig H., Kemler R.;
RT "Plakoglobin and beta-catenin: distinct but closely related.";
RL Science 257:1142-1144 (1992).
RN [2]
RP REVISIONS TO 294 AND 296.
RA Butz S.;
RL "FUNCTION: COMMON JUNCTIONAL PLAQUE PROTEIN. THE MEMBRANE-ASSOCIATED FORM. BELONGS TO THE BETA-CATENIN FAMILY. ASSOCIATED PLAQUES ARE ARCHITECTURAL ELEMENTS IN AN IMPORTANT STRATEGIC POSITION TO INFLUENCE THE ARRANGEMENT AND FUNCTION OF BOTH THE CYTOSKELETON AND THE CELLS WITHIN THE TISSUE. THE PRESENCE OF PLAQUEGLOBIN IN BOTH THE DESMOSOMES AND IN THE INTERMEDIATE JUNCTIONS SUGGESTS THAT IT PLAYS A CENTRAL ROLE IN THE STRUCTURE AND FUNCTION OF SUBMEMBRANOUS PLAQUES."
RN [2]
RP SUBUNIT: HOMODIMER.
RN [2]
RP SUBCELLULAR LOCATION: CYTOPLASMIC IN A SOLUBLE AND MEMBRANE-ASSOCIATED FORM.
RN [2]
RP SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
RN [2]
RP SIMILARITY: CONTAINS AT LEAST 9 ARM REPEATS.
RN [2]
RP This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

PROSITE, PS50176; ARM_REPEAT; 9.
Developmental protein; Segmentation polarity protein; Repeat;
Cell adhesion; Cytoskeleton; Structural protein.
DOMAIN 1 148 ASP/GLU-RICH (ACIDIC).
REPEAT 139 178 ARM 1.
REPEAT 180 221 ARM 2.
REPEAT 223 262 ARM 3.
REPEAT 265 304 ARM 4.
REPEAT 349 387 ARM 5.
REPEAT 388 427 ARM 6.
REPEAT 430 471 ARM 7.
REPEAT 477 517 ARM 8.
REPEAT 587 626 ARM 9.
REPEAT 628 667 ARM 10.
SEQUENCE 813 AA; 88237 MW; 73E61F59BDBFA580 CRC64;

Query Match 74.2%; Score 144; DB 1; Length 813;
Best Local Similarity 74.4%; Pred. No. 2.3e-11;
Matches 29; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 AATKQEGMGLLTGLVQLGSDDDINVTCAAGILSLNLT 39
DB 389 AATKYGLEPLQLSLVQLVQLASTDNNVTCAAGILSLNLT 427
RESULT 8
PLAK_XENLA
ID PLAK_XENLA STANDARD; PRT; 738 AA.
AC P30998;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Junction plakoglobin (Desmoplakin III).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93093332; Pubmed=1459359;
RA Fouquet B., Zimbelmann R., Franke W.W.;
RT "Identification of plakoglobin in oocytes and early embryos of Xenopus laevis: maternal expression of a gene encoding a junctional plaque protein.";
RL Differentiation 51:187-194 (1992).
RN [2]
RP SEQUENCE OF 133-292 FROM N.A.
RX MEDLINE=93012479; Pubmed=1397690;
RA de Marais A.A., Moon R.T.;
RT "The armadillo homologs beta-catenin and plakoglobin are differentially expressed during early development of Xenopus laevis.";
RL Dev. Biol. 153:337-346 (1992).
RN [2]
RP FUNCTION: COMMON JUNCTIONAL PLAQUE PROTEIN. THE MEMBRANE-ASSOCIATED PLAQUES ARE ARCHITECTURAL ELEMENTS IN AN IMPORTANT STRATEGIC POSITION TO INFLUENCE THE ARRANGEMENT AND FUNCTION OF BOTH THE CYTOSKELETON AND THE CELLS WITHIN THE TISSUE. THE PRESENCE OF PLAQUEGLOBIN IN BOTH THE DESMOSOMES AND IN THE INTERMEDIATE JUNCTIONS SUGGESTS THAT IT PLAYS A CENTRAL ROLE IN THE STRUCTURE AND FUNCTION OF SUBMEMBRANOUS PLAQUES."
RN [2]
RP SUBUNIT: HOMODIMER.
RN [2]
RP SUBCELLULAR LOCATION: CYTOPLASMIC IN A SOLUBLE AND MEMBRANE-ASSOCIATED FORM.
RN [2]
RP SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
RN [2]
RP SIMILARITY: CONTAINS 10 ARM REPEATS.
RN [2]
RP This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

entitles requires a license agreement (See <http://www.isb-sib.ch/announcement> or send an email to license@isb-sib.ch).

CC EMBL: M95593; AAA49931.1; -
CC EMBL: X67078; CAA47463.1; -
CC PIR: S24636; S24636.
CC PIR: S35093; S35093.
CC HSSP: Q02248; 2BCT.
CC InterPro: IPR000225; Armadillo.
CC Pfam: PF00514; Armadillo_seg; 11.
CC SMART: SM00185; ARM; 10.
CC PROSITE: PS0176; ARM_REPEAT; 9.
CC Cell Adhesion: Cytoskeleton; Structural protein; Repeat.
FT REPEAT 128 167 ARM 1.
FT REPEAT 212 251 ARM 2.
FT REPEAT 254 293 ARM 3.
FT REPEAT 338 377 ARM 4.
FT REPEAT 379 416 ARM 5.
FT REPEAT 419 460 ARM 6.
FT REPEAT 466 506 ARM 7.
FT REPEAT 508 547 ARM 8.
FT REPEAT 570 609 ARM 9.
FT REPEAT 611 650 ARM 10.
FT CONFLICT 185 185 I -> V (IN REF. 2).
FT CONFLICT 226 226 A -> T (IN REF. 2).
SQ SEQUENCE 738 AA; 81711 MW; 569DBE69D08BBC58 CRC64;

Query Match 67.5%; Score 131; DB 1; Length 738;
Best Local Similarity 65.8%; Pred. No. 1.e-09;
Matches 25; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 ATKQEGMGLLTGLVQLGSDDDINVTCAAGILSLNLT 39
DB 379 ATKQGLDNLVQLGSDDDINVTCAAGILSLNLT 416
RESULT 9
PLAK_MOUSE
ID PLAK_MOUSE STANDARD; PRT; 621 AA.
AC Q02257;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Junction plakoglobin (Desmoplakin III) (Fragment).
GN JUP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93376536; Pubmed=1509266;
RA Butz S., Stappert J., Weissig H., Kemler R.;
RT "Plakoglobin and beta-catenin: distinct but closely related.";
RL Science 257:1142-1144 (1992).
RN [2]
RP REVISIONS TO 294 AND 296.
RA Butz S.;
RL "FUNCTION: COMMON JUNCTIONAL PLAQUE PROTEIN. THE MEMBRANE-ASSOCIATED PLAQUES ARE ARCHITECTURAL ELEMENTS IN AN IMPORTANT STRATEGIC POSITION TO INFLUENCE THE ARRANGEMENT AND FUNCTION OF BOTH THE CYTOSKELETON AND THE CELLS WITHIN THE TISSUE. THE PRESENCE OF PLAKOGLOBIN IN BOTH THE DESMOSOMES AND IN THE INTERMEDIATE JUNCTIONS SUGGESTS THAT IT PLAYS A CENTRAL ROLE IN THE STRUCTURE AND FUNCTION OF SUBMEMBRANOUS PLAQUES."
CC -!- SUBUNIT: HOMODIMER.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC IN A SOLUBLE AND MEMBRANE-ASSOCIATED FORM.
CC -!- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -!- SIMILARITY: CONTAINS AT LEAST 9 ARM REPEATS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

DR PROSITE, PS50176; ARM_REPEAT; 9.
DR Developmental protein; Segmentation polarity protein; Repeat;
DR Cell adhesion; Cytoskeleton; Structural protein.
DR DOMAIN 1 148 ASP/GLU-RICH (ACIDIC).
FT REPEAT 139 178 ARM 1.
FT REPEAT 180 221 ARM 2.
FT REPEAT 223 262 ARM 3.
FT REPEAT 265 304 ARM 4.
FT REPEAT 349 387 ARM 5.
FT REPEAT 388 427 ARM 6.
FT REPEAT 430 471 ARM 7.
FT REPEAT 477 517 ARM 8.
FT REPEAT 587 626 ARM 9.
FT REPEAT 628 667 ARM 10.
SQ SEQUENCE 813 AA; 88237 MW; 73E61F59BDBFA580 CRC64;

Query Match 74.2%; Score 144; DB 1; Length 813;
Best Local Similarity 74.4%; Pred. No. 2.3e-11;
Matches 29; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 AATKQEGMGLLTGLVQLGSDDDINVTCAAGILSLNLT 39
DB 389 AATKYGLEPLQLSLVQLVQLASTDVNVVTCAAGILSLNLT 427
RESULT 8
PLAK_XENLA
ID PLAK_XENLA STANDARD; PRT; 738 AA.
AC P30998;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Junction plakoglobin (Desmoplakin III).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93093332; Pubmed=1459359;
RA Fouquet B., Zimbelmann R., Franke W.W.;
RT "Identification of plakoglobin in oocytes and early embryos of Xenopus laevis: maternal expression of a gene encoding a junctional plaque protein.";
RL Differentiation 51:187-194 (1992).
RN [2]
RP SEQUENCE OF 133-292 FROM N.A.
RX MEDLINE=93012479; Pubmed=1397690;
RA de Marais A.A., Moon R.T.;
RT "The armadillo homologs beta-catenin and plakoglobin are differentially expressed during early development of Xenopus laevis.";
RL Dev. Biol. 153:337-346 (1992).
CC -!- FUNCTION: COMMON JUNCTIONAL PLAQUE PROTEIN. THE MEMBRANE-ASSOCIATED PLAQUES ARE ARCHITECTURAL ELEMENTS IN AN IMPORTANT STRATEGIC POSITION TO INFLUENCE THE ARRANGEMENT AND FUNCTION OF BOTH THE CYTOSKELETON AND THE CELLS WITHIN THE TISSUE. THE PRESENCE OF PLAKOGLOBIN IN BOTH THE DESMOSOMES AND IN THE INTERMEDIATE JUNCTIONS SUGGESTS THAT IT PLAYS A CENTRAL ROLE IN THE STRUCTURE AND FUNCTION OF SUBMEMBRANOUS PLAQUES."
CC -!- SUBUNIT: HOMODIMER.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC IN A SOLUBLE AND MEMBRANE-ASSOCIATED FORM.
CC -!- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -!- SIMILARITY: CONTAINS 10 ARM REPEATS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; M90365; AAB02885.1; -
 DR PIR; S35092; S35092.
 DR HSSP; Q02248; 1DOW.
 DR MGD; MGI:96650; Jup.
 DR InterPro; IPR000225; Armadillo.
 DR Pfam; PF00514; Armadillo_seg; 11.
 DR SMART; SM00185; ARM; 8.
 DR PROSITE; PS0176; ARM_REPEAT; 9.
 KW Cell adhesion; Cytoskeleton; Structural protein; Repeat.
 FT NON_TER 1 1
 FT REPEAT 8 47 ARM 1.
 FT REPEAT 92 131 ARM 2.
 FT REPEAT 134 173 ARM 3.
 FT REPEAT 218 257 ARM 4.
 FT REPEAT 259 296 ARM 5.
 FT REPEAT 299 340 ARM 6.
 FT REPEAT 346 386 ARM 7.
 FT REPEAT 388 427 ARM 8.
 FT REPEAT 450 489 ARM 9.
 SQ SEQUENCE 621 AA; 68111 MW; 17CF444607422BAA CRC64;

Query Match 66.0%; Score 128; DB 1; Length 621;
 Best Local Similarity 65.8%; Pred. No. 2.3e-09;
 Matches 25; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

OY 2 ATKQSGMGLLTGLVOLLGSDINVTCAAGTSLNLT 39
 DB 259 ATKQSGLESLVILYNQLNSVDVNVLTCTGTLNLT 296

RESULT 10
 PLAK_HUMAN
 ID PLAK_HUMAN STANDARD; PRT; 743 AA.
 AC P14923;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Junction plakoglobin (Desmoplakin III).
 GN JUP OR DP3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-69264555; PubMed=2726765;
 RX Franke W.W., Goldschmidt M.D., Zimbelmann R., Mueller H.M.,
 RA Schiller D.L., Cowin P.;
 RT "Molecular cloning and amino acid sequence of human plakoglobin, the
 RT common junctional plaque protein.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:4027-4031(1989).
 CC -!- FUNCTION: COMMON JUNCTIONAL PLAQUE PROTEIN. THE MEMBRANE-
 CC ASSOCIATED PLAQUES ARE ARCHITECTURAL ELEMENTS IN AN IMPORTANT
 CC STRATEGIC POSITION TO INFLUENCE THE ARRANGEMENT AND FUNCTION
 CC OF BOTH THE CYTOSKELETON AND THE CELLS WITHIN THE TISSUE. THE
 CC PRESENCE OF PLAQUEGLIBIN AND BOTH THE DESMOSOMES AND IN THE
 CC INTERMEDIATE JUNCTIONS SUGGESTS THAT IT PLAYS A CENTRAL ROLE IN
 CC THE STRUCTURE AND FUNCTION OF SUBMEMBRANOUS PLAQUES.
 CC -!- SUBUNIT: HOMODIMER.
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC IN A SOLUBLE AND MEMBRANE-
 CC ASSOCIATED FORM.
 CC -!- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
 CC -!- SIMILARITY: CONTAINS 8 ARM REPEATS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; M23410; AAA64895.1; -
 DR PIR; A32905; A32905.
 DR HSSP; Q02248; 2BCT.
 DR MIM; M173325;
 DR InterPro; IPR000225; Armadillo.
 DR Pfam; PF00514; Armadillo_seg; 10.
 DR SMART; SM00185; ARM; 8.
 DR PROSITE; PS0176; ARM_REPEAT; 8.
 KW Cell adhesion; Cytoskeleton; Structural protein; Repeat.
 FT INIT_MET 0 0
 FT REPEAT 141 179 ARM 1.
 FT REPEAT 183 224 ARM 2.
 FT REPEAT 225 252 ARM 3.
 FT REPEAT 308 351 ARM 4.
 FT REPEAT 389 430 ARM 5.
 FT REPEAT 431 473 ARM 6.
 FT REPEAT 478 521 ARM 7.
 FT REPEAT 582 624 ARM 8.
 FT CONFLICT 96 99 GOL A (IN REF. 1; AAA64895).
 FT CONFLICT 139 139 V -> A (IN REF. 1; AAA64895).
 SQ SEQUENCE 743 AA; 81498 MW; 472741F400D388FD CRC64;

Query Match 66.0%; Score 128; DB 1; Length 743;
 Best Local Similarity 65.8%; Pred. No. 2.7e-09;
 Matches 25; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

OY 2 ATKQSGMGLLTGLVOLLGSDINVTCAAGTSLNLT 39
 DB 381 ATKQSGLESLVILYNQLNSVDVNVLTCTGTLNLT 418

RESULT 11
 CTNB_TRIGR
 ID CTNB_TRIGR STANDARD; PRT; 820 AA.
 AC P35223;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Beta-catenin.
 OS Tripneustes gratilla (Hawaiian sea urchin).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoidea; Euechinoidea; Echinacea; Temnopleuroidea; Toxopneustidae;
 OC Tripneustes.
 OX NCBI_TaxID=7673;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-93305730; PubMed=8318544;
 RX Rosenthal E.T.;
 RT "Identification of homologues to beta-catenin/plakoglobin/armadillo
 RT in two invertebrates, Urechis caupo and Tripneustes gratilla.";
 RL Biochim. Biophys. Acta 1173:337-341(1993).
 CC -!- FUNCTION: BINDS TO THE CYTOPLASMIC DOMAIN OF THE CELL-CELL
 CC ADHESION MOLECULE E-CADHERIN, AND PERHAPS TO OTHER (MEMBRANE)
 CC PROTEINS. THE ASSOCIATION OF CATENINS TO CADHERINS PRODUCES A
 CC COMPLEX WHICH IS LINKED TO THE ACTIN FILAMENT NETWORK AND WHICH
 CC SEEMS TO BE OF PRIMARY IMPORTANCE FOR CADHERINS CELL-ADHESION
 CC PROPERTIES.
 CC -!- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
 CC -!- SIMILARITY: CONTAINS 10 ARM REPEATS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: L10354; AAA30089.1; -
 DR PIR: S33794; S33794.
 DR HSSP: Q02248; IDOW.
 DR InterPro: IPR000225; Armadillo.
 DR Pfam: PF00514; Armadillo_seg; 12.
 DR SMART: SM00185; ARM; 12.
 DR PROSITE: PS00176; ARM_REPEAT; 9.
 KW Cell adhesion: Cytoskeleton; Structural protein; Repeat.
 FT REPEAT 157 196 ARM 1.
 FT REPEAT 199 239 ARM 2.
 FT REPEAT 241 280 ARM 3.
 FT REPEAT 283 322 ARM 4.
 FT REPEAT 367 405 ARM 5.
 FT REPEAT 406 445 ARM 6.
 FT REPEAT 448 489 ARM 7.
 FT REPEAT 495 535 ARM 8.
 FT REPEAT 603 642 ARM 9.
 FT REPEAT 644 683 ARM 10.
 SQ SEQUENCE 820 AA; 89361 MW; 57255E0F57795FD3 CRC64;

Query Match 61.9%; Score 120; DB 1; Length 820;
 Best Local Similarity 73.5%; Pred. No. 34e-08;
 Matches 25; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 6 EGMEGLLTGLVQLGSDINNVTCAGILSLNLC 39
 DB 412 DTEPLQLVQLGSDINNVTCAGILSLNLC 445

RESULT 12
 Y787_SVNY3
 ID Y787_SVNY3 STANDARD; PRT; 322 AA.
 AC Q55946;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 34.7 kDa protein sll0787.
 GN sll0787.
 OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=1148;
 RN SEQUENCE FROM N.A.
 RP MEDLINE=96127529; PubMed=8590279;
 RX Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
 RA Sugita M., Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
 region from map positions 64% to 94% of the genome.";
 RC DNA Res. 2:153-166(1995).
 CL -1- SIMILARITY: TO M.JANNASCHII MJ0640 AND MJ0799.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: D64005; BAA10714.1; -
 DR InterPro: IPR000728; AIRS_related.
 DR Pfam: PF00586; AIRS; 1.
 DR Pfam: PF02769; AIRS_C; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 322 AA; 34723 MW; C6B434C5B3C0460D CRC64;

Query Match 28.9%; Score 56; DB 1; Length 322;
 Best Local Similarity 31.6%; Pred. No. 3.7;
 Matches 12; Conservative 9; Mismatches 11; Indels 6; Gaps 1;

QY 2 ATKQEGMEGLLTGLVQLGSD-----DINVTCAGI 33
 DB 210 AKDVSVMGGIIGTALMLETSCGAILDLDAISCPAGL 247

RESULT 13
 AC02_SVNY3
 ID AC02_SVNY3 STANDARD; PRT; 868 AA.
 AC P74582;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Aconitate hydratase 2 (EC 4.2.1.3) (Citrate hydro-lyase 2)
 DE (Aconitase 2).
 GN ACNB OR SUR0665.
 OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=1148;
 RN [1];
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hirotsawa M., Sugita M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,
 RA Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,
 RA Yanada M., Yasuda M., Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 Synechocystis sp. strain PCC6803. II. Sequence determination of the
 entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 CC -1- CATALYTIC ACTIVITY: Citrate = cis-aconitate + H(2)O.
 CC -1- PATHWAY: TRICARBOXYLIC ACID CYCLE.
 CC -1- MISCELLANEOUS: ACONITASE HAS AN ACTIVE (4FE-4S) AND AN INACTIVE
 CC (3FE-4S) FORMS. THE ACTIVE (4FE-4S) CLUSTER IS PART OF THE
 CC CATALYTIC SITE THAT INTERCONVERTS CITRATE, CIS-ACONITASE, AND
 CC ISOCITRATE.

-1- SIMILARITY: BELONGS TO THE ACONITASE/IPM ISOMERASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: D90916; BAA18689.1; -
 DR HSSP: PI7297; IDHY.
 DR InterPro: IPR001030; Aconitase.
 DR Pfam: PF00330; aconitase; 2.
 DR ProDom: PD000511; Aconitase; 1.
 DR PROSITE: PS00450; ACONITASE_1; 1.
 DR PROSITE: PS01244; ACONITASE_2; FALSE NEG.
 KW Lyase; tricarboxylic acid cycle; Iron-sulfur; 4Fe-4S;
 KW Complete proteome.
 FT METAL 700 IRON (IRON-SULFUR CLUSTER)
 FT METAL 700 (BY SIMILARITY).
 FT METAL 758 IRON (IRON-SULFUR CLUSTER)
 FT METAL 758 (BY SIMILARITY).
 FT METAL 761 IRON (IRON-SULFUR CLUSTER)
 FT METAL 761 (BY SIMILARITY).
 SQ SEQUENCE 868 AA; 93551 MW; EBF93B0F966A94EE CRC64;

Query Match 28.1%; Score 54.5; DB 1; Length 868;
 Best Local Similarity 38.5%; Pred. No. 15;
 Matches 15; Conservative 5; Mismatches 10; Indels 9; Gaps 1;

QY 6 EGMEGLLT-----LVQLGSDINNVTCAGILS 35

Query Match 27.8%; Score 54; DB 1; Length 525;
Best Local Similarity 39.1%; Pred. No. 11;
Matches 1 9; Conservative 7; Mismatches 7; Indels

QY . 10 GLLGTLVQLIGSDDINVTCAAG 32
.: : : : : : : : : :
Db . 462 GVIGRYGRILGDNDINIATMQVG 484

AC	P73821;				
DT	15-JUL-1998	(Rel. 36, Created)			
DT	15-JUL-1998	(Rel. 36, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			

GN	SERA OR SLL1908.
OS	Synechocystis sp. (strain PCC 6803).
OC	Bacteria; Cyanobacteria; Chroococcales; Synechocystis.

[1], "SEQUENCE FROM N.A.",
RX MEDLINE=97061201; Pubmed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Na

RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.
RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M.,
RA Tabata S.;

CC -!- CATALYTIC ACTIVITY: 3-phosphoglycerate + NAD(+) = 3-DNA Res. 3:109-136(1996).

CC OF L-SERINE BIOSYNTHESIS.
CC -|- PATHWAY: FIRST COMMITTED STEP IN THE "PHOSPHORYLATED
CC -|- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDRO
CC DEHYDROGENASES FAMILY

CC This SWISS-PROT entry is copyright. It is produced through
CC between the Swiss Institute of Bioinformatics and the
CC the European Bioinformatics Institute. There are no re-

CC
CC
CC
CC
CC

modified and this statement is not removed. Usage by
entities requires a license agreement (See <http://www.isib-sib.ch>).
or send an email to license@isb-sib.ch.

CC

DR InterPro; IPR002162; D_2_hydroxyacid_DH.
DR Pfam; PF00389; 2-Hacid_DH; 1.
DR Pfam; PF02826; 2-Hacid_DH_C; 1.
DR Pfam; PF01412; 2-Hacid_DH_C; 1.

DR PROSITE; PS00670; D_2_HYDROXYACID_DH_2; 1.
DR PROSITE; PS00671; D_2_HYDROXYACID_DH_3; 1.
KW Serine biosynthesis; Oxidoreductase; NAD; Complete prote

FT	ACT_SITE	287	BY SIMILARITY.
FT	ACT_SITE	305	BY SIMILARITY.
SQ	SEQUENCE	554 AA; 59221 MW;	1D25DCDE7A8DBD19 CRC64;

Query, Match 27.3%; Score 53; DB 1: Length 55
 Best, Local Similarity 33.3%;
 Matches 4; 10; Conservative 7; Mismatches 13; Indels

QY . . . 3 TRQEGEGLGTLVQLLGSDLLNVVTCAG 32
| | : : | : : : |
Db 485 TLHRDMPGIIGKIGSLGSEFNVIASMQVG 514

```
Query_Match      ...27.8%; Score 54; DB 1; Length 525;
Best Local Similarity 39.1%; Pred. No. 11;
Matches! 9; Conservative 7; Mismatches 7; Indels 0; Gaps 0;
```

QY 10 GLLGTLVQLIGSDDINNVVTCAAG 32
 | : | : : | : | : | : | : |
Db 462 GVIGRVGRILGDNDINIATMOVG 484

AC	P73821;	Created
DT	15-JUL-1998	(Rel. 36)
DT	15-JUL-1998	Last sequence update)
DT	Rel. 36	Last sequence update)
DT	15-OCT-2001	(Rel. 40)
DT	16-OCT-2001	Last annotation update)

GN SERA OR SLL1908.
OS *Synechocystis* sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; *Synechocystis*.

[1] SEQUENCE FROM N.A. . . .
RN MEDLINE=97061201; PubMed=8905231;
RP KANEKO T., Sato S., Kotani H., Tanaka A.,
RX Asamizu E., Nakamura Y.,
RA

RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpō S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S., Analysis of the genome of the unicellular cyanobacterium
RA *Nostoc muscorum* by using a bacterial artificial chromosome library

RT
CC
CL

CC -1- PATHWAY: FIRST COMMITTED STEP IN THE "PHOSPHORYLATED" PATHWAY.
CC OF L-SERINE BIOSYNTHESIS.
CC OF L-SERINE BIOSYNTHESIS.
CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
CC DEHYDROGENASES FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or 'send an email to license@isb-sib.ch).
CC -----
CC

DR InterPro: IPR002162; D_2_hydroxyacid_DH.
DR Pfam: PF00389; 2-Hacid_DH; 1.
DR Pfam: PF02826; 2-Hacid_DH_C; 1.
DR Pfam: PF03040; 2-Hacid_DH_C; 1.
DR Pfam: PF03040; 2-Hacid_DH_C; 1.

DR PROSITE; PS00670; D-2-HYDROXYACID_DH_2; 1.
DR PROSITE; PS00671; D-2-HYDROXYACID_DH_3; 1.
KW Serine biosynthesis; Oxidoreductase; NAD; Complete proteome.

FT	ACT_SITE	287	BY SIMILARITY.
FT	ACT_SITE	305.	BY SIMILARITY.
SQ	SEQUENCE	554 AA;	59221 MW; 1D25DCDE7A8DB1D9 CRC64;

Query: Watch 27.3%; Score 53; DB 1; Length 554;
 Best Local Similarity 33.3%; Pred. No. 16;
 Matches 10; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

QY . . . 3 TKQEGMEGLGLTVQLGSDLDINNVVICAAG 32
| | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | |
485 TLHRDMPGIIGKIGSLGSLGFNVNIAASMQVG 514
Db

Tue Jul 30 08:15:24 2002

Search completed: July 29, 2002, 16:09:55
Job time: 760 sec

us-09-641-104a-10.rsp

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 16:12:07 ; Search time 124.4 seconds
(without alignments)
54.235 Million cell updates/sec

Title: US-09-641-104A-10

Perfect score: 194

Sequence: 1 AATKQEGMEGLLGLTLLVQLLGSDINVTCAAGILSNLTC 39

Scoring table: BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	194	100.0	780	13 Q90424	Q90424 brachydanio
2	194	100.0	781	11 Q9D335	Q9D335 mus musculus
3	194	100.0	781	13 Q42486	Q42486 gallus gall
4	150	77.3	806	5 Q25100	Q25100 hydra magni
5	145	74.7	729	13 Q9PVF7	Q9PVF7 brachydanio
6	144	74.2	769	5 Q9NLA4	Q9NLA4 ciona intes
7	144	74.2	773	5 Q76152	Q76152 ciona savig
8	128	66.0	745	4 Q15151	Q15151 homo sapien
9	128	66.0	745	4 Q9BWC4	Q9BWC4 homo sapien
10	126	64.9	745	11 P70565	P70565 rattus norv
11	120	61.9	821	5 Q61229	Q61229 lytechinus
12	71	36.6	678	5 Q44326	Q44326 caenorhabdi
13	64	33.0	796	5 Q10953	Q10953 caenorhabdi
14	62	32.0	1299	10 Q49300	Q49300 arabidopsis
15	59	30.4	103	3 Q01596	Q01596 pneumocysti
16	59	30.4	3021	12 Q92933	Q92933 hepatitis c

17 59 30.4 3021 12 Q81258 hepatitis c
18 57 29.4 228 5 Q9VM18 drosophila
19 57 29.4 3021 12 Q68870 hepatitis c
20 55.5 28.6 248 16 Q9A920 caulobacter
21 55 28.4 550 3 Q43028 schizosacch
22 55 28.4 2931 5 Q9W2C6 drosophila
23 55 28.4 3023 12 Q81487 hepatitis c
24 54 27.8 419 3 Q9UT06 schizosacch
25 54 27.8 458 10 Q49659 arabidopsis
26 54 27.8 3021 12 Q81495 hepatitis c
27 53.5 27.6 526 17 Q9HMX1 halobacteri
28 53 27.3 166 16 Q31769 bacillus su
29 53 27.3 373 16 Q92R63 rhizobium m
30 53 27.3 936 2 Q9F809 paenibacill
31 52.5 27.1 333 11 Q9QZ03 mus musculu
32 52 26.8 555 10 Q9F158 arabidopsis
33 52 26.8 555 10 Q94BNS arabidopsis
34 52 26.8 556 10 Q9LF97 arabidopsis
35 52 26.8 606 5 Q9VWF6 drosophila
36 52 26.8 664 3 Q9P7W7 schizosacch
37 52 26.8 706 10 Q9FYA4 arabidopsis
38 52 26.8 800 10 Q9FTH3 oryza sativ
39 52 26.8 2062 5 Q9SPH6 dictyostell
40 52 26.8 3013 12 Q92530 hepatitis c
41 52 26.8 3019 12 Q68801 hepatitis c
42 51.5 26.5 465 5 Q9V905 drosophila
43 51.5 26.5 514 8 Q98RP2 quillardia
44 51.5 26.5 561 2 Q52072 rhodococcus
45 51 26.3 111 1 Q9P9K7 methanosarc

ALIGNMENTS

RESULT 1

ID Q90424 PRELIMINARY; PRT; 780 AA.
AC Q90424;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE B-CATENIN.
GN CTNNB.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96122902; PubMed=8562427;
RA Kelly G.M., Erezylmaz D.F., Moon R.T.;
RT "Induction of a secondary embryonic axis in zebrafish occurs following the overexpression of beta-catenin.";
RL Mech. Dev. 53:261-273(1995).
DR EMRL; U41081; AAC59732.1; .
DR HSSP; P35222; 1G3T
DR ZFIN; ZDB-GENE-980526-362; cttnb.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 12.
DR SMART; SM00185; ARM; 11.
DR PROSITE; PS0176; ARM-REPEAT; 8.
SQ SEQUENCE 780 AA; 85542 MW; D7A1FB80F94066DC CRC64;

Query Match 100.0%; Score 194; DB 13; Length 780;
Best Local Similarity 100.0%; Pred. No. 4.6e-18;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATKQEGMEGLLGLTLLVQLLGSDINVTCAAGILSNLTC 39

DB 390 AATKQEGMEGLLGLTLLVQLLGSDINVTCAAGILSNLTC 428

```

RESULT 2
Q9D335 ID Q9D335 PRELIMINARY; PRT; 781 AA.
AC Q9D335;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ADULT MALE COLON CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,
DE CLONE:9030417H18, FULL INSERT SEQUENCE.
GN CATNB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=COLON;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa T., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohetsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
DR EMBL; AK018515; BAB31250.1; -
DR HSSP; P35222; IG3J.
DR MGD; MGI:88276; Catnb.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 12.
DR SMART; SM00185; ARM; 12.
DR PROSITE; PS0176; ARM_REPEAT; 7.
SQ SEQUENCE 781 AA; 85546 MW; 937538C3B5CD75D1 CRC64;

Query Match 100.0%; Score 194; DB 11; Length 781;
Best Local Similarity 100.0%; Pred. No. 4.6e-18;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATKQEGMEGLLTGLVQLLGSDDINVTCAAGILSNLTC 39
|||||
Db 391 AATKQEGMEGLLTGLVQLLGSDDINVTCAAGILSNLTC 429
|||||

RESULT 3
O42486 ID O42486 PRELIMINARY; PRT; 781 AA.
AC O42486;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE BETA CATENIN.
GN CHC9AT.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN
[1]

```

```

SEQUENCE FROM N.A.
RC STRAIN=WHITE LEHORN; TISSUE=DORSAL SKIN;
RX MEDLINE=97464068; PubMed=9322759;
RA Liu J., Chuong C.M., Widelitz R.B.;
RT Isolation and characterization of chicken beta-catenin.;
RL Gene 196:201-207 (1997).
DR EMBL; U82964; AAB08056.1; -
DR HSSP; P35222; IG3J.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 12.
DR SMART; SM00185; ARM; 11.
DR PROSITE; PS50176; ARM_REPEAT; 8.
SQ SEQUENCE 781 AA; 85438 MW; 6D205D9A4DBAC562 CRC64;

Query Match 100.0%; Score 194; DB 13; Length 781;
Best Local Similarity 100.0%; Pred. No. 4.6e-18;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATKQEGMEGLLTGLVQLLGSDDINVTCAAGILSNLTC 39
|||||
Db 391 AATKQEGMEGLLTGLVQLLGSDDINVTCAAGILSNLTC 429
|||||

RESULT 4
Q25100 ID Q25100 PRELIMINARY; PRT; 806 AA.
AC Q25100;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE BETA-CATENIN.
OS Hydra magnipapillata (Hydra).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydrozoa; Anthomedusae;
OC Hydridae; Hydra.
OX NCBI_TaxID=6085;
RN
[1]
SEQUENCE FROM N.A.
RC STRAIN=WILDTYPE 105;
RX MEDLINE=96257271; PubMed=8654977;
RA Hobmayer E., Hatta M., Fischer R., Fujisawa T., Holstein T.W.,
RA Sugiyama T.;
RT Identification of a Hydra homologue of the beta-
catenin/plakoglobin/armadillo gene family.;
RL Gene 172:155-159 (1996).
DR EMBL; U36781; AAC47137.1; -
DR HSSP; Q02248; 2BCT.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 10.
DR SMART; SM00185; ARM; 10.
DR PROSITE; PS50176; ARM_REPEAT; 7.
SQ SEQUENCE 806 AA; 90462 MW; 689E5E982CD5051A CRC64;

Query Match 77.3%; Score 150; DB 5; Length 806;
Best Local Similarity 76.3%; Pred. No. 5.1e-12;
Matches 29; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 2 ATKQEGMEGLLTGLVQLLGSDDINVTCAAGILSNLTC 39
|||||
Db 451 ATKQEGMEGLLTGLVQLLGSDDINVTCAAGILSNLTC 488
|||||

RESULT 5
Q9PVF7 ID Q9PVF7 PRELIMINARY; PRT; 729 AA.
AC Q9PVF7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CELL-ADHESION PROTEIN PLAKOGLOBIN.
GN JUP.
OS Brachydanio rerio (zebrafish) (zebra danio).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 OC Cypriniformes; Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99386700; PubMed=10456847;
 RA Cerda J., Reidenbach S., Pratzel S., Franke W.W.;
 RT "Caderin-catenin complexes during zebrafish oogenesis: heterotypic
 junctions between oocytes and follicle cells.";
 RL Biol. Reprod. 61:692-704(1999).
 DR EMBL; AF099738; AAD56592.1; -.
 DR HSSP; Q02248; IDOW.
 DR ZFIN; ZDB-GENE-991207-22; jup.
 DR InterPro; IPR000225; Armadillo.
 DR Pfam; PF00514; Armadillo_seg; 11.
 DR SMART; SM00185; ARM; 8
 DR PROSITE; PS50176; ARM_REPEAT; 8.
 SQ SEQUENCE 729 AA; 80033 MW; 91E00417B4FD8CEE CRC64;

Query Match 74.7%; Score 145; DB 13; Length 729;
 Best Local Similarity 71.8%; Pred. No. 2.2e-11;
 Matches 28; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 1 AATKQEGMEGLLTGLVQLLGSDDDINVTCAAGILSNLTC 39
 DB 371 AATKQDGMENLLQVLGLSADDINMLTCATGVLNLTC 409
 ||||:|||| || || || ||||:||||:|||||||

RESULT 6
 Q9NL44 PRELIMINARY; PRT; 769 AA.
 ID Q9NL44;
 AC Q9NL44;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE BETA-CATENIN.
 GN CIBETA-CATENIN.
 OS Ciona intestinalis.
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;
 OC Cionidae; Ciona.
 OX NCBI_TaxID=7719;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Imai K., Takada N., Satoh N., Satou Y.;
 RT "An essential role of beta-catenin in the endoderm specification of
 ascidian embryo.";
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB031543; BAA92185.1; -.
 DR HSSP; P35222; 1G3J.
 DR InterPro; IPR000225; Armadillo.
 DR Pfam; PF00514; Armadillo_seg; 11.
 DR SMART; SM00185; ARM; 11.
 DR PROSITE; PS50176; ARM_REPEAT; 6.
 SQ SEQUENCE 769 AA; 84703 MW; F61CC489B436E1BC CRC64;

Query Match 74.2%; Score 144; DB 5; Length 769;
 Best Local Similarity 76.9%; Pred. No. 3.2e-11;
 Matches 30; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 1 AATKQEGMEGLLTGLVQLLGSDDDINVTCAAGILSNLTC 39
 DB 381 AGTKQDHVENLQMLVQLLSSNDINVTCAAGILSNLTC 419
 ||||:|||| || || || ||||:||||:|||||||

RESULT 7
 O76152 PRELIMINARY; PRT; 773 AA.
 ID O76152;
 AC O76152;
 DT 01-NOV-1998 (TReMBLrel. 08, Created)
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)

DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE BETA-CATENIN.
 OS Ciona savignyi.
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;
 OC Cionidae; Ciona.
 OX NCBI_TaxID=51511;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98443204; PubMed=9769178;
 RA Yoshida S., Marikawa Y., Satoh N.;
 RT "Regulation of the trunk-tail patterning in the ascidian embryo: a
 possible interaction of cascades between lithium/beta-catenin and
 localized maternal factor pem.";
 RL Dev. Biol. 202:264-279(1998).
 DR EMBL; AB012160; BAA32789.1; -.
 DR HSSP; P35222; 1G3J.
 DR InterPro; IPR000225; Armadillo.
 DR Pfam; PF00514; Armadillo_seg; 11.
 DR SMART; SM00185; ARM; 9.
 DR PROSITE; PS50176; ARM_REPEAT; 7.
 SQ SEQUENCE 773 AA; 85217 MW; C1340CF82AFEBDAB CRC64;

Query Match 74.2%; Score 144; DB 5; Length 773;
 Best Local Similarity 76.9%; Pred. No. 3.2e-11;
 Matches 30; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 1 AATKQEGMEGLLTGLVQLLGSDDDINVTCAAGILSNLTC 39
 DB 380 AGTKQDQVENLQMLVQLLSSNDINVTCAAGILSNLTC 418
 ||||:|||| || || || ||||:||||:|||||||

RESULT 8
 Q15151 PRELIMINARY; PRT; 745 AA.
 ID Q15151;
 AC Q15151; Q15093;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE PLAKOGLOBIN.
 GN JUP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89264555; PubMed=2726765;
 RA Franke W.W., Goldschmidt M.D., Zimbelmann R., Mueller H.M.,
 RA Schiller D.L., Cowin P.;
 RT "Molecular cloning and amino acid sequence of human plakoglobin, the
 common junctional plaque protein.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:4027-4031(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Zimbelmann R.;
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 239-409 FROM N.A.
 RC TISSUE=PLACENTA;
 RX MEDLINE=96157724; PubMed=8576101;
 RA Ozawa M., Nukui K., Toyoyama H., Ohi Y.;
 RT "Cloning of an alternative form of plakoglobin (gamma-catenin) lacking
 the fourth armadillo repeat.";
 RL J. Biochem. 118:836-840(1995).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX PubMed=11016852;
 RA Whitlock N.V., Eady R.A.J., McGrath J.A.;
 RT "Genomic Organization and amplification of the human plakoglobin
 gene.";
 RL Exp. Dermatol. 9:323-326(2000).
 DR EMBL; Z68228; CAA92522.1; -.

```

DR ENBL; D50808; BAA09435.1; -.
DR ENBL; AF306723; AAG16727.1; -.
DR ENBL; AF333862; AAG16727.1; JOINED.
DR HSSP; Q02248; 2BCT.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 11.
DR SMART; SM00185; ARM; 9.
DR PROSITE; PS50176; ARM_REPEAT; 7.
DR SEQUENCE; 745 AA; 81744 MW; 3519A0973748BCF4 CRC64;
SQ

```

Query Match 66.0%; Score 128; DB 4; Length 745;
Best Local Similarity 65.8%; Pred. No. 4.9e-09;
Matches 25; Conservative 4; Mismatches 9; Indels

Qy 2 ATKQEGMEGLLTGLVQLGSDDINVVTCAGILSNLTC 39
 |||||: : | | :|||:| | |||||
 Db 383 ATKQEGLSVKILVNQLSDVDNVLTCAATGILSNLTC 420

RESULT	9
Q9BWC4	
ID	Q9BWC4
AC	PRELIMINARY;
CC	PRT: 745 AA.
DT	01-JUN-2001 (TReMBLrel. 17, Created)
DT	01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT	01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE	DE JUNCTION PLAKOGLOBIN.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;

RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=LUNG CARCINOMA;	
RA	Traustberg R.;	
RL	Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.	

[2]
 RN SEQUENCE FROM N.A.
 RP TISSUE=PLACENTA, CHORIOCARCINOMA;
 RC Strausberg R.;
 RA Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RL

```

DR ENBL; BC000441; AAH00441.1; -
DR ENBL; BC011865; AAH11865.1; -
DR HSSP; Q02248; 2BCT
DR Interpro; ipR00225; Armadillo
DR Pfam; PF00514; Armadillo_seg; 11.
DR SMART; SM00185; ARM; 12
DR PROSITE; P500176; ARM_REPEAT; 7
DR PROSITE; 745_AA; 81736 MW; 34DF7BF84718BCF4
DR SO SOURCE

```

Query Match 66.0%; Score 128; DB 4; Length 745;
Best Local Similarity 65.8%; Pred. No. 4.9e-09;
Matches 25; Conservative 4; Mismatches 9; Indels

QY 2 ATKQEGMEGLLTGLVQLLGSDDINVVTCAGILSNLTC 39
 | | | | | : | | | | | : | | | | |
pB 383 ATKOEGLSVLKLYNOLSVDDNVLTCA TGTLNLTC 420

RESULT	10	
ID	P70565	PRELIMINARY; PRT; 745 AA.
AC	P70565;	
DT	01-FEB-1997	(T-EMBLrel. 02, Created)
DT	01-FEB-1997	(T-EMBLrel. 02, Last sequence update)
DT	01-DEC-2001	(T-EMBLrel. 19, Last annotation update)
DE	PLAKOGLOBIN.	
OS	Rattus norvegicus (Rat).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus;	
OX	NCBI_TaxID=10116;	

RN
RP
RC
RD
RE
RF
RG
RH
RI
RJ
RK
RL
RM
RN
RO
RP
RQ
RR
RS
RT
RU
RV
RW
RX
RY
RZ

[1]...
SEQUENCE FROM N.A.
STRAIN-SPRAGUE-DAWLEY; TISSUE-VAGINA;
Umegita Y., Liao S.;
RA Molecular cloning and sequencing of the rat plakoglobin cDNA."
RT Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.

RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE-VAGINA;
RA Hiipakka R.A.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.

DR	EMBL:	U58858; AAB06317.1; -
DR	HSSP:	Q02248; 2BCT.

```

DR      .interpro;
DR      .interpro; IPR000225; Armadillo.
DR      pfam; pf00514; Armadillo seg: 11.

```

```
DR      1100014; NAME=SM00185;
DR      SMART; SM00185; ARM; 8.
DR      PROSITE: PS50176; ARM REPEAT. 7
```

```
DR PROSITE, PS00170, KRM_REPEAT, 7.  
SQ SEQUENCE 745 AA; 81777 MW; 9E2F52910A7ACD41 CRC64;
```

Query Match	64.9%;	Score 126;	DB 11;	Length 745;
Best Local Similarity	65.8%;	Pred. No. 9.2e-09;		
Matches 25; Conservative	4;	Mismatches 9;	Indels	

QY : 2 ATKQEGMEGLLTGLVQLLGSDDINNVTCAGIISNLT 39
 |||||: : || :||:|||| : |||||
 Db 383 ATKQEGLENVLKILNOLSVDDNVLTCAATHLSNLT 420

RESULT	ID	PRELIMINARY;	PRT;	821 AA.
11	061229			
	AC	061229;		
	DT	01-AUG-1998 (TReMBLrel. 07, Created)		
	DT	01-AUG-1998 (TReMBLrel. 07, Last sequence update)		
	DT	01-JUN-2001 (TReMBLrel. 17, Last annotation update)		
	DE	BETA-CATENIN.		
	OS	Lytechinus variegatus (sea urchin).		
	OC	Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;		
	OC	Echinozoa; Euechinozoa; Echinacea; Temnopleurozoa; Toxopneustidae;		
	OC	Lytechinus.		
	OX	NCEI_TaxID=7654;		

RN [1].
 RP SEQUENCE FROM N. A.
 RP MEDLINE=98104237; PubMed=9441670;
 RX Miller J.R., McClay D.R.;
 RA "Changes in the pattern of adherens junction-associated beta-catenin
 RT accompany morphogenesis in the sea urchin embryo.";
 RT dev. Biol. 192:310-322(1997).
 RT

```

DR EMEL; U34814; AAC06340.1; -
DR HSP; Q07248; IDOW
DR InterPro; IPR002086; Aldehyde_dehydr.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 12.
DR SMART; SM00185; ARM; 12.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
DR PROSITE; PS0176; ARM_REPEAT; 8.
DR PROSITE; 821.AA; 89558 MW; 71E21D562A9C5AD CRC64.
DR SO

```

Query Match	61.9%	Score 120;	DB 5;	Length 821;
Best Local Similarity	73.5%	Pred. No. 6.8e-08;		
Matches: 25;	Conservative	3;	Mismatches 6;	Indels 0;
				Gaps 0;

QY .6 EGMGLLTQLLGGDDINVVTCAGILSLNLC 39
 : : | | | | | : | | | | |
 Db 413 DDIEPLLOMLVOLLASNDINVVTCAGILSLNLC 446

RESULT. 12
O44326
ID O44326
AC O44326;
PRELIMINARY;
PRT; 678 AA.

RC STRAIN-C.B-17 SCID/MOUSE-DERIVED PNEUMOCYSTIS CARINII;
 RX MEDLINE-94222508; PubMed-8168913;
 RA Wright T.W., Simpson-Haidaris P.J., Gigliotti F., Harmsen A.G.,
 RA Haidaris C.G.;
 RT "Conserved sequence homology of cysteine-rich regions in genes
 encoding glycoprotein A in Pneumocystis carinii derived from different
 RT host species";
 RL Infect. Immun. 62:1513-1519(1994).
 DR EMBL; U05262; AAA18864.1; -;
 FT NON_TER 1
 FT NON_TER 103
 SQ SEQUENCE 103 AA; 12128 MW; B7AA05EE826438B5 CRC64;

Query Match 30.4%; Score 59; DB 3; Length 103;
 Best Local Similarity 38.2%; Pred. No. 1.6;
 Matches 13; Conservative 6; Mismatches 11; Indels 4; Gaps 1;

QY 5 QEGMEGLGLTLVQLLGSDDINVTCAAGILSNLT 38
 |||:|:|:|
 Db 18 QEGLGLIGN---LGLADENLEECQKSVVGNVT 47

Search completed: July 29, 2002, 16:12:09
 Job time: 479 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 16:07:56 ; Search time 158.47 seconds
(without alignments)
30.840 Million cell updates/sec

Title: US-09-641-104a-11
Perfect score: 227
Sequence: 1 NNYKNKMVMQVGGIEALVR.....AGDREITEPAICALRHLS 44

Scoring table: BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 747574 seqs, 11107396 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : A_Geneseq_032802.*
- 1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
 - 2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
 - 3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
 - 4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
 - 5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
 - 6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
 - 7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
 - 8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
 - 9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
 - 10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
 - 11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
 - 12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
 - 13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
 - 14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
 - 15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
 - 16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
 - 17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
 - 18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
 - 19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
 - 20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
 - 21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
 - 22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	227	100.0	44	AAV33222	Human beta-catenin
2	227	100.0	781	AA807290	Human beta-catenin
3	227	100.0	781	AAU70740	Human beta-catenin
4	227	100.0	781	AAU28118	Novel human secret
5	227	100.0	781	AAE06038	Human beta-catenin
6	227	100.0	781	AAE06039	Mouse beta-catenin
7	225	99.1	44	AAV33237	Human beta-catenin
8	187	82.4	840	AB860196	Drosophila melanog
9	187	82.4	840	AB865819	Drosophila melanog
10	187	82.4	840	AB865821	Drosophila melanog
11	166	73.1	345	21 AAB56810	Human prostate can

12	4166	73.1	700	12	AA11354	Cellular Receptor
13	6119	52.4	800	22	AAU28306	Novel human secret
14	58.5	25.8	797	21	AAB36462	Human plakophilin-
15	58.5	25.8	797	22	AAU79226	Human protein SEQ
16	58.5	25.8	817	22	AAU80210	Human protein SEQ
17	57.5	25.3	531	22	ABG00802	Novel human diagno
18	57	25.1	160	21	AAG14043	Arabidopsis thalia
19	57	25.1	161	21	AAG39419	Arabidopsis thalia
20	57	25.1	164	21	AAG14042	Arabidopsis thalia
21	57	25.1	165	21	AAG39418	Arabidopsis thalia
22	57	25.1	178	21	AAG14041	Arabidopsis thalia
23	57	25.1	206	21	AAG39417	Arabidopsis thalia
24	55.5	24.4	223	21	AAU32318	Corn beta-carotene
25	54.5	24.0	978	22	AAU33885	Staphylococcus aur
26	54.5	24.0	987	22	AAU36599	Staphylococcus aur
27	54	23.8	1106	22	AAU03552	Human protein kina
28	52.5	23.1	939	22	AAE10797	Human catenin p120
29	52	22.9	461	21	AAG36942	Arabidopsis thalia
30	52	22.9	480	21	AAG36941	Arabidopsis thalia
31	52	22.9	564	21	AAG36940	Arabidopsis thalia
32	51.5	22.7	201	22	AAU40279	Propionibacterium
33	51.5	22.7	267	22	AAU09551	Human polypeptide
34	51.5	22.7	797	21	AAB36466	Mouse plakophilin-
35	51	22.5	191	21	AAU74676	Neisseria meningit
36	51	22.5	775	22	ABB60574	Drosophila melanog
37	51	22.5	1009	22	ABG20820	Novel human diagno
38	51	22.5	1110	22	AAE03648	Human extracellular
39	51	22.5	1192	20	AAU23899	Human resenilin bl
40	51	22.5	1211	18	AAU24560	Presenilin-interac
41	51	22.5	1211	21	AAU92336	Human p0071. Homo
42	51	22.5	1445	12	AAU12108	N-terminal deleted
43	51	22.5	1522	10	AAU93357	Sequence of the ca
44	51	22.5	1644	18	AAU13504	B. bronchiseptica
45	51	22.5	1645	18	AAU13502	B. pertussis adeny

ALIGNMENTS

RESULT 1
AAV33222
ID AAV33222 standard; peptide: 44 AA.
XX
AC AAV33222;
XX
DT 18-NOV-1999 (first entry)
XX
DE Human beta-catenin protein armadillo repeat arm8 fragment.
DE
DE Beta-catenin; human; armadillo repeat; treatment; human disease; LCF-1;
KW modulator; interaction domain; transcription factor; TCF-4; oncogenic;
KW tumor suppressor gene; conductin; E-cadherin; screening; carcinoma; APC;
KW Wnt signalling pathway; tumor development; anti-oncogenic; melanoma;
KW organ regeneration; tissue regeneration; hair growth.
XX
OS Homo sapiens.
XX
PN DE19909251-AL.
XX
PD 26-AUG-1999.
XX
PF 22-FEB-1999; 99DE-1009251.
XX
PR 21-FEB-1998; 98DE-1007390.
XX
PA (DELB-) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.
PI Bitchmeier W, Von Kries J;
XX
DR WPI; 1999-470389/40.
XX
PT Agents for treating human diseases, particularly cancer, modulate
interaction of beta-catenin with transcription factors or tumor

PT suppressor gene products -

XX
PS Disclosure; Page 7; 16pp; German.

XX This invention describes a novel agent (A) for treating human disease
CC which is based on substances (I) that modulate (inhibit or promote) the
CC interaction of beta-catenin with transcription factors or products of
CC tumor suppressor genes. The invention also describes (a) peptides (II)
CC comprising part of the LEF-1/TCF-4 transcription factors, or their
CC variants and mutants; (b) peptides and related molecules (III) from the
CC armadillo domain (arm units 3-8) of beta-catenin, and mutants of the
CC complete beta-catenin molecule, that include at least one of the
CC specific interaction domains for LEF-1, TCF-4, APC, conductin or
CC E-cadherin; and (c) enzyme-linked immunosorbent assay (ELISA) for
CC screening substance libraries for compounds that modulate interaction of
CC beta-catenin with LEF-1/TCF, APC, conductin or E-cadherin. Beta-catenin
CC is a key compound in the Wnt signalling pathway and is involved in
CC development of tumors. Generally its interaction with LEF-1 or TCF-4 is
CC oncogenic but interaction with APC, conductin or E-cadherin is
CC anti-oncogenic. (A) which inhibit interaction are particularly used to
CC treat tumors, especially carcinoma of the colon and melanoma, but also,
CC where they promote interaction, to stimulate regeneration of organs and
CC tissues, specifically hair growth. AAY3217-13322 represent human
CC beta-catenin armadillo repeat fragments described in the method of the
CC invention.

XX Sequence 44 AA;

Query Match 100.0%; Score 227; DB 20; Length 44;

Best Local Similarity 100.0%; Pred. No. 3.4e-26;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNYKNMVCQGGIEALVTVLRAGDREDITEPAICALRHLS 44
DB 1 nnyknmvcqggiealvtrvragdreditepaicalrhls 44

RESULT 2

AAB07290
ID AAB07290 standard; Protein; 781 AA.

XX AAB07290;

XX 01-NOV-2000 (first entry)

XX Human beta catenin.

XX Beta catenin; cadherin; metastasis; cadherin-associated protein;
KW human; colorectal cancer; melanoma; antisense oligonucleotide;
KW gene therapy.

XX Homo sapiens.

XX US6066500-A.

XX 23-MAY-2000.

XX 25-JUN-1999; 99US-0344519.

XX 25-JUN-1999; 99US-0344519.

XX (ISIS-) ISIS PHARM INC.

XX Bennett CF, Cowser LM;

XX WPI: 2000-410651/35.

XX N-PSDB: AAA58320.

XX New antisense compounds targeting nucleic acids encoding human beta
PT catenin (HBC) useful for treating diseases associated with HBC
PT expression and as prophylaxis to prevent or delay infection,
PT inflammation or tumor formation

XX

PS Example 13; Columns 45-52; 35pp; English.

XX Beta catenin is a member of the catenin family of cytosolic proteins and
CC a key member of the Wnt signalling pathway. Catenins interact with the
CC cytoplasmic domains of cadherin glycoproteins, and are important in
CC maintaining cell adhesiveness. The loss of cell adhesiveness is
CC implicated in metastasis. Beta catenin is also known as
CC cadherin-associated protein and is implicated in colorectal cancer and
CC melanoma. The present sequence is the human beta catenin protein. The
CC coding sequence of this protein was used in the present invention to
CC design antisense oligonucleotides (AA58327-A58366). The
CC oligonucleotides are capable of hybridizing to human beta catenin, in
CC order to inhibit expression of human beta catenin. The oligonucleotides
CC may be used in gene therapy for colorectal cancer or melanoma.

XX Sequence 781 AA;

Query Match 100.0%; Score 227; DB 21; Length 781;

Best Local Similarity 100.0%; Pred. No. 9.6e-25;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNYKNMVCQGGIEALVTVLRAGDREDITEPAICALRHLS 44

DB 430 nnyknmvcqggiealvtrvragdreditepaicalrhls 473

RESULT 3

AAY70740

ID AAY70740 standard; Protein; 781 AA.

XX AAY70740;

XX 24-JUL-2000 (first entry)

XX Human beta-catenin.

XX Human beta-catenin; Wnt antagonist; contraceptive; contraceptive vaccine;
KW oocyte development; female primate contraception; oocyte viability;
KW monoclonal antibody; Wnt signalling.

XX Homo sapiens.

XX WO200021555-A1.

XX 20-APR-2000.

XX 13-OCT-1999; 99WO-US23640.

XX 15-OCT-1998; 98US-0104355.

XX (HARD;) HARVARD COLLEGE.

XX McMahon AP, Parr BA, Vaino S;

XX WPI: 2000-317845/27.

XX Contraceptive composition for inhibiting oocyte development in a female
PT primate comprises a Wnt polypeptide antagonist

XX Example 3; Page 26; 57pp; English.

XX The patent discloses a method of female primate contraception comprising
CC administering an antagonist of a Wnt polypeptide, inhibiting oocyte
CC development. Wnt polypeptides are useful for promotive maturation of an
CC immature oocyte. Wnt polypeptides are also useful for increasing the
CC number of mature oocytes and to enhance oocyte viability. Soluble
CC fragments of Wnt polypeptides have the ability to inhibit Wnt signalling,
CC e.g., by blocking binding of a naturally-occurring Wnt protein to its
CC receptor. They may be used to generate monoclonal antibodies which can
CC inhibit oocyte development. The present sequence is the human
CC beta-catenin protein. Dominant negative mutants of beta-catenin lack one

CC or more armadillo-like repeats which participate in cadherin binding.
 CC Other mutants include those lacking amino acids 555-781 or 424-781,
 CC or 1-422 of beta-catenin.

XX Sequence 781 AA;

Query Match 100.0%; Score 227; DB 21; Length 781;
 Best Local Similarity 100.0%; Pred. No. 9.6e-25;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNYKNKMMVCQVGIEALVRLVLRAGDREDITEPAICALRHLS 44
 |||||
 Db 430 nnyknmmvcvggiealvrlvragdreditepaicalrhls 473

RESULT 4

AAU28118
 ID AAU28118 standard; Protein; 781 AA.

XX AAU28118;

XX 18-DEC-2001 (first entry)

XX Novel human secretory protein, Seq ID No 287.

XX Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
 KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
 KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
 KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
 KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
 KW gut protection; lung; liver fibrosis; immune deficiency; infection;
 KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
 KW fertility; analgesic; pain; antigen.

XX Homo sapiens.

XX WO200166689-A2.

XX 13-SEP-2001.

XX 05-MAR-2001; 2001WO-US04942.

XX 07-MAR-2000; 2000US-0519705.

XX 19-MAY-2000; 2000US-0574454.

XX 17-JUN-2000; 2000US-0596193.

XX 14-JUL-2000; 2000US-0616847.

XX 19-SEP-2000; 2000US-0665363.

XX 20-OCT-2000; 2000US-0693267.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
 PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
 XX N-PSDB; AAS45018.

XX WPI; 2001-589934/66.

XX Novel polypeptides and nucleic acids obtained from cDNA libraries
 PT prepared from various human tissues, for diagnosis and treatment of
 PT cancer, neurological, inflammatory, and autoimmune disorders -

XX Example 4; SEQ ID No 287; 107pp; English.

XX The invention relates to novel isolated human secreted polypeptides (I)
 CC and polynucleotides (II). (I) and (II) are useful for treating
 CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
 CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
 CC involved in increasing haematopoiesis, stem cell survival, bone growth
 CC and remodelling. (I), (II) and modulators of (II) are useful for
 CC prophylaxis or treatment of one or more cancers. (II) is also useful for
 CC creating transgenic animals useful for studying the in vivo activities of

CC the polypeptide as well as for studying modulators of the polypeptides.
 CC (I) induces the proliferation of neural cells and regeneration of nerve
 CC and brain tissue and is useful for the treatment of central and
 CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
 CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
 CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
 CC activity, regulation of haematopoiesis and is useful for treating myeloid
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 CC tissue growth, and in tissue repair, healing of burns, incisions,
 CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
 CC disorders, or periodontal disease. Furthermore, (I) is also useful for
 CC gut protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues, various immune deficiencies and
 CC disorders including severe combined immunodeficiency (SCID), bacterial or
 CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
 CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
 CC reactions and conditions, such as asthma or other respiratory problems.
 CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
 CC fertility, metabolism, catabolism, anabolism, storage or elimination of
 CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
 CC analgesic effects or other pain reducing effects, immunoglobulin like
 CC activity and can act as an antigen in a vaccine composition to raise an
 CC immune response. AAU28020-AAU28395 represent novel human secreted protein
 CC amino acid sequences of the invention.

XX Sequence 781 AA;

Query Match 100.0%; Score 227; DB 22; Length 781;
 Best Local Similarity 100.0%; Pred. No. 9.6e-25;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNYKNKMMVCQVGIEALVRLVLRAGDREDITEPAICALRHLS 44
 |||||
 Db 430 nnyknmmvcvggiealvrlvragdreditepaicalrhls 473

RESULT 5

AAE06038
 ID AAE06038 standard; Protein; 781 AA.

XX AAE06038;

XX 25-SEP-2001 (first entry)

XX Human beta-catenin protein #1.

XX Human; stem cell culturing; progenitor cell; pluripotential phenotype;
 KW transplantation; haematopoietic function; allogeneic recipient;
 KW signalling pathway; beta-catenin.

XX Homo sapiens.

XX WO200152649-A1.

XX 26-JUL-2001.

XX 17-JAN-2001; 2001WO-US01459.

XX 18-JAN-2000; 2000US-0176786.

XX (STPRD) UNIV LELAND STANFORD JUNIOR.

XX Reya T, Nusse R, Weissman IL;

XX WPI; 2001-465328/50.

XX N-PSDB; AAD11164.

XX In vitro expansion of mammalian stem or progenitor cells, useful for
 PT producing cells that retain their pluripotential phenotype after
 PT expansion for use in transplantation, involves increasing the levels of
 PT beta-catenin in the cell -

XX Disclosure; Page 23-24; 33pp; English.

XX The present invention relates to a method for in vitro expansion of

CC mammalian stem or progenitor cells, which comprising increasing the

CC intracellular concentration of beta-catenin in a progenitor or stem cell

CC in an in vitro culture medium for a period sufficient for the progenitor

CC or stem cell to divide. The number of cells having the functional

CC phenotype of the stem or progenitor cells is expanded. The method is

CC useful for culturing stem cells and progenitor cells, which retain their

CC pluripotential phenotype after expansion, in vitro. The expanded cell

CC populations are useful as a source of stem cells, e.g. to reconstitute

CC function in a host that is deficient in a particular cell lineage or

CC lineages. The expanded cell populations are also useful in

CC transplantation to restore haematopoietic function to autologous or

CC allogeneic recipients. The present sequence is human beta-catenin

CC protein. Beta-catenin is a pivotal player in the signalling pathway

CC initiated by Wnt proteins, which are mediators of several developmental

CC processes.

XX Sequence 781 AA;

SQ

Query Match 100.0%; Score 227; DB 22; Length 781;

Best Local Similarity 100.0%; Pred. No. 9.6e-25;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNYKNKMVCQGGIEALVTVLRAGDREDITEPAICALRHLS 44

Db 430 nnyknkmvcqggiealvrtviragdreiditepaicalrhls 473

RESULT 6

AAE06039

ID AAE06039 standard; Protein; 781 AA.

XX

AC AAE06039;

XX

DT 25-SEP-2001 (first entry)

XX

DE Mouse beta-catenin protein #2.

XX

XX Mouse; stem cell culturing; progenitor cell; pluripotential phenotype;

KW transplantation; haematopoietic function; allogeneic recipient;

KW signalling pathway; beta-catenin.

XX

OS Mus musculus.

XX

PN WO200152649-A1.

XX

PD 26-JUL-2001.

XX

PF 17-JAN-2001; 2001WO-US01459.

XX

PR 18-JAN-2000; 2000US-0176786.

XX

PA (STRD) UNIV LELAND STANFORD JUNIOR.

XX

PI Reya T, Nusse R, Weissman IL;

XX

DR WPI; 2001-465328/50.

XX

DR N-PSDB; AAD11165.

XX

PT In vitro expansion of mammalian stem or progenitor cells, useful for

PT producing cells that retain their pluripotential phenotype after

PT expansion for use in transplantation, involves increasing the levels of

PT beta-catenin in the cell.

XX

XX Disclosure; Page 28-30; 33pp; English.

XX The present invention relates to a method for in vitro expansion of

CC mammalian stem or progenitor cells, which comprising increasing the

CC intracellular concentration of beta-catenin in a progenitor or stem cell

CC in an in vitro culture medium for a period sufficient for the progenitor

CC or stem cell to divide. The number of cells having the functional

CC phenotype of the stem or progenitor cells is expanded. The method is

CC useful for culturing stem cells and progenitor cells, which retain their

CC pluripotential phenotype after expansion, in vitro. The expanded cell

CC populations are useful as a source of stem cells, e.g. to reconstitute

CC function in a host that is deficient in a particular cell lineage or

CC lineages. The expanded cell populations are also useful in

CC transplantation to restore haematopoietic function to autologous or

CC allogeneic recipients. The present sequence is mouse beta-catenin

CC protein. Beta-catenin is a pivotal player in the signalling pathway

CC initiated by Wnt proteins, which are mediators of several developmental

CC processes.

XX Sequence 781 AA;

SQ

Query Match 100.0%; Score 227; DB 22; Length 781;

Best Local Similarity 100.0%; Pred. No. 9.6e-25;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNYKNKMVCQGGIEALVTVLRAGDREDITEPAICALRHLS 44

Db 430 nnyknkmvcqggiealvrtviragdreiditepaicalrhls 473

RESULT 7

AAV33237

ID AAV33237 standard; peptide; 44 AA.

XX

AC AAV33237;

XX

DT 18-NOV-1999 (first entry)

XX

DE Human beta-catenin protein mutant armadillo repeat arm 8.

XX

KW Beta-catenin; human; armadillo repeat; treatment; human disease; LEF-1;

KW modulator; interaction domain; transcription factor; TCF-4; oncogenic;

KW tumor suppressor gene; conductin; E-cadherin; screening; carcinoma; APC;

KW Wnt signalling pathway; tumor development; anti-oncogenic; melanoma;

KW organ regeneration; tissue regeneration; hair growth; mutant.

XX

OS Homo sapiens.

OS Synthetic.

XX

PN DE19909251-A1.

XX

PD 26-AUG-1999.

XX

PF 22-FEB-1999; 99DE-1009251.

XX

PR 21-FEB-1998; 98DE-1007390.

XX

PA (DELB-) DELBRUECK-CENT MOLEKULARE MEDIZIN MAX.

XX

PI Birchmeier W, Von Kries J;

XX

DR WPI; 1999-470389/40.

XX

XX Agents for treating human diseases, particularly cancer, modulate

PT interaction of beta-catenin with transcription factors or tumor

PT suppressor gene products

XX

XX Example 4; Fig 5; 16pp; German.

XX This invention describes a novel agent (A) for treating human disease

CC which is based on substances (I) that modulate (inhibit or promote) the

CC interaction of beta-catenin with transcription factors or products of

CC tumor suppressor genes. The invention also describes (a) peptides (II)

CC comprising part of the LEF-1/TCF-4 transcription factors, or their

CC variants and mutants; (b) peptides and related molecules (III) from the

CC armadillo domain (arm units 3-8) of beta-catenin, and mutants of the

CC complete beta-catenin molecule, that include at least one of the

CC specific interaction domains for LEF-1, TCF-4, APC, conductin or
 CC E-cadherin; and (c) enzyme-linked immunosorbent assay (ELISA) for
 CC screening substance libraries for compounds that modulate interaction of
 CC beta-catenin with LEF-1/TCF, APC, conductin or E-cadherin. Beta-catenin
 CC is a key compound in the Wnt signalling pathway and is involved in
 CC development of tumors. Generally its interaction with LEF-1 or TCF-4 is
 CC oncogenic but interaction with APC, conductin or E-cadherin is
 CC anti-oncogenic. (A) which inhibit interaction are particularly used to
 CC treat tumors, especially carcinoma of the colon and melanoma, but also,
 CC where they promote interaction, to stimulate regeneration of organs and
 CC tissues, specifically hair growth. AAY3230-Y33241 represent mutant
 CC human beta-catenin armadillo repeat fragments described in the method of
 CC the invention.

XX Sequence 44 AA;

Query Match 99.1%; Score 225; DB 20; Length 44;
 Best Local Similarity 97.7%; Pred. No. 6.7e-26;
 Matches 43; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NNYKNMVMVCQVGIEALVRTVLRAAGDREDITEPAICALRHLS 44
 |||||
 Db 1 nnyknmvmvcvggiealvrtvlragdreditepaicalrhls 44

RESULT 8

ABB60196
 ID ABB60196 standard; Protein; 840 AA.

XX ABB60196;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 7380.

XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL04299.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

XX Disclosure; SEQ ID NO 7380; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).

XX The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO

CC at:ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 840 AA;

Query Match 82.4%; Score 187; DB 22; Length 840;
 Best Local Similarity 75.0%; Pred. No. 8.6e-19;
 Matches 33; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

OY 1 NNYKNMVMVCQVGIEALVRTVLRAAGDREDITEPAICALRHLS 44
 |||||
 Db 435 nnyknmvmvcvggiealvrtvlragdreditepaicalrhls 478

RESULT 9

ABB65819

ID ABB65819 standard; Protein; 840 AA.

XX ABB65819;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 24249.

XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE') PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL09922.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

XX Disclosure; SEQ ID NO 24249; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).

XX The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at:ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 840 AA;

Query Match 82.4%; Score 187; DB 22; Length 840;

Best Local Similarity 75.0%; Pred. No. 8.6e-19;

Matches 33; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

OY 1 NNYKNMVMVCQVGIEALVRTVLRAAGDREDITEPAICALRHLS 44

Db 435 nnyknmvmvcvggiealvrtvlragdreditepaicalrhls 478

RESULT 10

ABB56821
ID AAB56821 standard; Protein; 840 AA.

XX
AC ABB56821;

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 24255.

XX Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL09924.

XX New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.

XX Disclosure; SEQ ID NO 24255; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072).

XX The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 840 AA;

Query Match

Best Local Similarity 82.4%; Score 187; DB 22; Length 840;

Matches 33; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 1 NNYKMKMVCQGGIEALVTVLRAGDREDITEPAICALRHLS 44

DB 435 nngnrkatvcqvggdalvrtiinagdrresitepavcalrhls 478

RESULT 11

AAB56810
ID AAB56810 standard; Protein; 345 AA.

XX AAB56810;

XX 13-MAR-2001 (first entry)

XX Human prostate cancer antigen protein sequence SEQ ID NO:1388.

KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
KW vulnery; gastrointestinal; nephrotropic; antiinfective; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW wound; infectious disease.

XX Homo sapiens.

XX WO200055174-A1

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05988.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (ROSE/) ROSEN C A.

XX Rosen CA, Ruben SM;

XX WPI; 2000-587513/35.

XX N-PSDB; AAF16013.

XX Prostate cancer associated gene sequences, referred to as prostate

XX cancer antigens, useful for treatment, prevention, and diagnosis of

XX disorders such as prostate cancer.

XX Claim 11; Page 1813-1815; 2338pp; English.

XX AAF15566 to AAF16505 encode the human prostate cancer associated

XX proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.

XX The prostate cancer antigens can have neuroprotective, cytostatic,

XX cardioactive, immunomodulatory, muscular, vulnery, gastrointestinal,

XX nephrotropic, antiinfective, gynaecological and antibacterial activities,

XX and can be used in gene therapy. The prostate cancer antigen

XX identification, as chromosome markers, and for numerous other diagnostic

XX or research purposes. The prostate cancer antigens may be used to treat

XX gastrointestinal, pulmonary, cardiovascular, renal, and proliferative

XX disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to

XX AAB57303 represent sequences used in the exemplification of the present

XX invention.

XX Sequence 345 AA;

Query Match 73.1%; Score 166; DB 21; Length 345;

Best Local Similarity 68.2%; Pred. No. 3.9e-16;

Matches 30; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 1 NNYKMKMVCQGGIEALVTVLRAGDREDITEPAICALRHLS 44

DB 21 nnskntlvqsgveallhailragdkdditepavcalrhls 64

RESULT 12

AAR11354

ID AAR11354 standard; Protein; 700 AA.

XX AAR11354;

XX 03-JUN-1991 (first entry)

XX Cellular Receptor 2 ligand.

XX Cellular receptor 2; CR2; binding site; BS; auto-immune disease;

XX Epstein-Barr Virus; EBV; B lymphocyte; ligand.

XX Synthetic.

FH	Key	Location/Qualifiers
FT	Binding-site	300..312
FT	Binding-site	/note= "fragment pref. included"
FT	Binding-site	303..309
FT	Binding-site	/note= "fragment most pref. included"
FT	Binding-site	385..400
FT	Binding-site	/note= "fragment pref. included"
FT	Binding-site	389..394
FT	Binding-site	/note= "fragment most pref. included"
XX		
WQ	WO103251-A.	
PX	21-MAR-1991.	
XX		
PD		
XX		
PF	04-SEP-1990;	90WO-US05027.
XX		
PR	20-APR-1990;	90US-0512118.
PR	08-SEP-1989;	89US-0404679.
XX	(CALB-) CALIF INST BIOLOGIC.	
PA	Lernhardt W;	
PI		
XX		
DR	WPI; 1991-101864/14.	
DR	N-PSDB; AAQ11140.	
XX		
PT	DNA segment encoding CR-2 ligand and CR2 binding site - used to	
PT	treat auto-immune disease, B-cell lymphoma and inhibit	
PT	Epstein-Barr virus infection	
XX		
PS	Disclosure; Fig 1; 129pp; English.	
XX		
CC	The ligand pref. includes a fragment indicated in the features,	
CC	or the total protein. The ligand pref. contains only a single BS	
CC	and has an amino acid sequence <100 (pref.<20) residues in length.	
CC	The ligand may also comprise a plurality of fragment 389..394.	
CC	A therapeutic compsn. contg. the polypeptide is used to stimulate	
CC	or inhibit B lymphocyte proliferation in patients with B cell	
CC	lymphoma. B lymphocytes and myeloma's can be stimulated in	
CC	patients with immunodeficiencies and immunoglobulin secretion by	
CC	hybridoma cultures can be boosted.	
CC	The compsn. can be administered to inhibit infection in vitro or in	
CC	vivo by Epstein-Barr Virus.	
CC	See also AAQ11140-42.	
XX		
SQ	Sequence 700 AA;	
	Query Match 73.1%; Score 166; DB 12; Length 700;	
	Best Local Similarity 68.2%; Pred. No. 8.9e-16;	
	Matches 30; Conservative 7; Mismatches 7; Indels 0; Gaps 0;	
QY	1 NNYKNKMYCVGGIEALVTVLRAGDREDITEPATCALRHLS 44	
Db	: + : : : :	
	-378 nnskntlvqtsgvgeallhailragdkdditepavcalrhls 421	
	RESULT 13	
AAU28306		
ID	AAU28306 standard; Protein; 800 AA.	
XX		
AC	AAU28306;	
XX		
DT	18-DEC-2001 (first entry)	
XX		
DE	Novel human secretory protein, Seq ID No 663.	
XX		
KW	Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;	
KW	ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;	
KW	transgenic animal; Alzheimer's disease; Parkinson's disease; burn;	
KW	amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;	
KW	ulcer; osteoporosis; bone degenerative disorder; periodontal disease;	
KW	gut protection; lung/liver fibrosis; immune deficiency; infection;	

severe combined immunodeficiency; SCID; autoimmune disorder; allergy; multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma; fertility; analgesic; pain; antigen.

Homo sapiens.

WO200166689-A2.

13-SEP-2001.

05-MAR-2001; 2001WO-US04942.

07-MAR-2000; 2000US-0519705.

19-MAY-2000; 2000US-0574454.

17-JUN-2000; 2000US-0596193.

14-JUL-2000; 2000US-0616847.

19-SEP-2000; 2000US-0665363.

20-OCT-2000; 2000US-0693267.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P; Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;

WPI: 2001-569934/66.

N-PSDB; AAS45206.

Novel polypeptides and nucleic acids obtained from cDNA libraries prepared from various human tissues, for diagnosis and treatment of cancer, neurological, inflammatory, and autoimmune disorders -

Example 2; SEQ ID No 663; 107pp; English.

The invention relates to novel isolated human secreted polypeptides (I) and polynucleotides (II). (I) and (II) are useful for treating inflammatory conditions such as arthritis, nephritis, Crohn's disease, ischaemia-reperfusion injury, shock, sepsis, immune responses, and is involved in increasing haematopoiesis, stem cell survival, bone growth and remodeling. (I), (II) and modulators of (II) are useful for prophylaxis or treatment of one or more cancers. (II) is also useful for creating transgenic animals useful for studying the in vivo activities of the polypeptide as well as for studying modulators of the polypeptides. (I) induces the proliferation of neural cells and regeneration of nerve and brain tissue and is useful for the treatment of central and peripheral nervous system diseases and neuropathies, such as Alzheimer's, Parkinson's disease, Huntington's disease, and amyotrophic lateral sclerosis. In addition, (I) is involved in chemotactic or chemokinetic activity, regulation of haematopoiesis and is useful for treating myeloid or lymphoid cell disorders, platelet disorders such as thrombocytopenia and for regeneration of bone, cartilage, tendon, ligament and/or nerve tissue growth, and in tissue repair, healing of burns, incisions, ulcers, for treating osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease. Furthermore, (I) is also useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, various immune deficiencies and disorders including severe combined immunodeficiency (SCID), bacterial or fungal infections, autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic reactions and conditions, such as asthma or other respiratory problems. In addition, (I) affects biorhythms or circadian cycles of rhythms, fertility, metabolism, catabolism, anabolism, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides analgesic effects or other pain reducing effects, immunoglobulin like activity and can act as an antigen in a vaccine composition to raise an immune response. AAU28020-AAU28395 represent novel human secreted protein amino acid sequences of the invention.

Sequence 800 AA;

Query-Match 52.4%; Score 119; DB 22; Length 800;
Best Local Similarity 63.0%; Pred. No. 9.4e-09;
Matches 29; Conservative 5; Mismatches 8; Indels 4; Gaps 3;

QY 1 NNYKNMVMVC-QVGGTEALVRT--VLRAGDR-EDITEPAICLRHL 42
 AAB36462
 DB 443 nnyknmmvcpkgygrrllrxlviragggktslepaicalrhl 488
 RESULT 14
 ID AAB36462 standard; Protein; 797 AA.
 XX AAB36462;
 DT 01-MAR-2001 (first entry)
 DE Human plakophilin-3 protein SEQ ID NO:2.
 XX Plakophilin-3; PKP3; catenin-like protein; characterisation; diagnosis;
 KW desmosome; epithelial cell; skin disease; dermatological; gene therapy;
 KW vaccine.
 OS Homo sapiens.
 XX
 PN WO200066619-A2.
 XX
 PD 09-NOV-2000.
 XX
 PF 28-APR-2000; 2000WO-EP04389.
 XX
 PR 30-APR-1999; 99EP-0870093.
 XX
 PA (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
 XX
 PI Van Roy F, Bonne S;
 XX
 DR WPI; 2000-687529/67.
 DR N-PSDB; AAC64792.
 XX
 PT Nucleic acids encoding Plakophilin-3 polypeptides, useful for treating
 PT skin diseases and disorders of epithelial tissue associated with
 PT inappropriate Plakophilin-3 expression and activity -
 XX
 PS Example 1; Fig 2; 132pp; English.
 XX
 CC The present invention describes an isolated or recombinant nucleic acid
 CC molecule (I) encoding a plakophilin-3 (PKP3), from humans, mice and
 CC Xenopus laevis (I) has dermatological activity, and can be used in
 CC gene therapy and for vaccines. (I) and the protein it encodes may be
 CC used in the prevention, treatment and diagnosis of diseases associated
 CC with inappropriate PKP3 expression, such as skin diseases and disorders
 CC affecting epithelial tissue. For example, (I) (and vectors containing
 CC (I)) and the PKP3 polypeptide may be used to treat disorders associated
 CC with decreased PKP3 expression by rectifying mutations or deletions in
 CC a patient's genome that affect the activity of PKP3 by expressing
 CC inactive proteins or to supplement the patients own production of PKP3
 CC polypeptides. Additionally, (I) may be used to produce PKP3, according
 CC to standard recombinant DNA methodology, by inserting the nucleic acids
 CC into a host cell and culturing the cell to express the protein. (I) and
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC assays to detect and quantitate the presence of similar nucleic acid
 CC sequences in samples, and hence which patients may be in need of
 CC restorative therapy. The PKP3 polypeptides may be used as antigens in
 CC the production of antibodies against PKP3 and in assays to identify
 CC modulators (agonists and antagonists) of PKP3 expression and activity.
 CC The anti-PKP3 antibodies and PKP3 antagonists may also be used to down
 CC regulate PKP3 expression and activity. The anti-PKP3 antibodies may also
 CC be used as diagnostic agents for detecting the presence of PKP3
 CC polypeptides in samples (e.g. by enzyme linked immunosorbant assay
 CC (ELISA)). PKP3 is a catenin-like protein, which is present in desmosomes
 CC and nuclei of epithelial cells. The present sequence represents human
 CC PKP3, as given in the present invention.
 XX
 SQ Sequence 797 AA;

Query Match 25.8%; Score 58.5; DB 21; Length 797;
 Best Local Similarity: 40.9%; Pred. No. 8.3;
 Matches: 18; Conservative 8; Mismatches 13; Indels 5; Gaps 3;
 QY 4 KNRMMVCQVGIEALVRTV--LRAGRED-ITEPAICLRHL 43
 DB 493 rkmrech-glvdalvtshaldagkcdksvenavcrlrnl 535
 RESULT 15
 ID AAM79226 standard; Protein; 797 AA.
 XX AAM79226;
 XX
 DT 06-NOV-2001 (first entry)
 DE Human protein SEQ ID NO 1888.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN WO200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US04098.
 XX
 PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0654936.
 PR 15-SEP-2000; 2000US-0663561.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xie AJ, Yang Y, Wejhrman T, Goodrich R;
 XX
 DR WPI; 2001-476283/51.
 DR N-PSDB; AAK52359.
 XX
 PT Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -
 XX
 PS Claim 20; Page 4278-4280; 6221pp; English.
 XX
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 XX
 SQ Sequence 797 AA;

Query Match 25.8%; Score 58.5; DB 22; Length 797;
 Best Local Similarity 40.9%; Pred. No. 8.3;
 Matches 18; Conservative 8; Mismatches 13; Indels 5; Gaps 3;

OY 4 KNKMMVCOVGGIEALYRTV---LRAGDRED-ITEPAICALRHILT 43
 Db 493 rqkmrech-glvdalvtsinhaidagkcdksvenavcvlrnlS 535

Search completed: July 29, 2002, 16:07:57
 Job time: 647 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 16:05:10 ; Search time 57.6 seconds
(without alignments)
18.658 Million cell updates/sec

Title: US-09-641-104A-11

Perfect score: 227

Sequence: 1 NNYKNMVCQVGIEALVR.....AGDREDITEPAICALRHLS 44

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/1aa/PTIUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	24.7	666	US-08-982-785A-11	Sequence 11, Appl
2	52.5	23.1	682	US-08-982-785A-9	Sequence 9, Appl
3	51.5	22.7	620	US-08-982-785A-10	Sequence 10, Appl
4	51	22.5	525	6 5183745-4	Patent No. 5183745
5	51	22.5	1489	6 5183745-2	Patent No. 5183745
6	51	22.5	1705	4 US-08-669-785-4	Sequence 4, Appl
7	51	22.5	1706	4 US-08-669-785-2	Sequence 2, Appl
8	51	22.5	1794	6 5183745-6	Patent No. 5183745
9	50.5	22.2	686	4 US-08-982-785A-8	Sequence 8, Appl
10	50.5	22.2	756	4 US-08-982-785A-2	Sequence 2, Appl
11	49	21.6	79	6 5284931-7	Patent No. 5284931
12	49	21.6	131	2 US-08-721-488-8	Sequence 8, Appl
13	49	21.6	185	6 5489533-4	Patent No. 5489533
14	49	21.6	262	2 US-08-719-697-10	Sequence 10, Appl
15	49	21.6	262	4 US-08-727-616A-10	Sequence 10, Appl
16	49	21.6	449	1 US-08-256-136-2	Sequence 2, Appl
17	49	21.6	449	2 US-08-950-737-2	Sequence 2, Appl
18	49	21.6	449	3 US-08-973-461A-8	Sequence 8, Appl
19	49	21.6	449	3 US-08-648-010-8	Sequence 8, Appl
20	49	21.6	453	6 5284931-3	Patent No. 5284931
21	49	21.6	480	2 US-08-425-989B-11	Sequence 11, Appl
22	49	21.6	505	3 US-08-318-039A-1	Sequence 1, Appl
23	49	21.6	505	3 US-08-318-038D-15	Sequence 15, Appl
24	49	21.6	505	4 US-08-227-496C-19	Sequence 19, Appl
25	49	21.6	505	4 US-08-435-568A-1	Sequence 1, Appl
26	49	21.6	505	4 US-09-240-915-6	Sequence 6, Appl
27	49	21.6	505	4 US-09-591-435-6	Sequence 6, Appl

28	49	21.6	507	1 US-08-464-083-14	Sequence 14, Appl
29	49	21.6	507	2 US-08-469-587A-14	Sequence 14, Appl
30	49	21.6	531	2 US-08-789-078-3	Sequence 3, Appl
31	49	21.6	531	2 US-08-752-633-3	Sequence 3, Appl
32	49	21.6	531	5 PCT-US95-04886-3	Sequence 3, Appl
33	49	21.6	532	1 US-07-618-286-1	Sequence 1, Appl
34	49	21.6	532	2 US-08-196-003-3	Sequence 3, Appl
35	49	21.6	532	2 US-08-483-389-118	Sequence 118, App
36	49	21.6	532	2 US-08-689-870-12	Sequence 12, Appl
37	49	21.6	532	3 US-08-933-824-3	Sequence 3, Appl
38	49	21.6	532	4 US-09-264-466-3	Sequence 3, Appl
39	49	21.6	532	4 US-09-009-490A-87	Sequence 87, Appl
40	49	21.6	532	6 5284931-2	Patent No. 5284931
41	49	21.6	680	4 US-08-227-496C-15	Sequence 15, Appl
42	49	21.6	740	5 PCT-US93-00031-17	Sequence 17, Appl
43	48	21.1	161	3 US-08-973-068-7	Sequence 3, Appl
44	47	20.7	44	4 US-08-982-785A-3	Sequence 2, Appl
45	47	20.7	1196	4 US-08-881-706-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-982-785A-11
: Sequence 11, Application US/08982785A
: Patent No. 6258929
: GENERAL INFORMATION:
: APPLICANT: Kosik, Kenneth S.
: APPLICANT: Zhou, Jianhua
: TITLE OF INVENTION: ALARM RELATED PEPTIDES AND
: NUCLEIC ACIDS AND DIAGNOSIS USING THEM
: NUMBER OF SEQUENCES: 13
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 225 Franklin Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02110-2804
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: OPERATING SYSTEM: Windows 95
: SOFTWARE: FastSeq for Windows Version 2.0b
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/982,785A
: FILING DATE: 02-DEC-1997
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/031,556
: FILING DATE: 02-DEC-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Freeman, John W.
: REGISTRATION NUMBER: 29,066
: REFERENCE/DOCKET NUMBER: 05311/018001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617/542-5070
: TELEFAX: 617/542-8906
: TELEX: 200154
: INFORMATION FOR SEQ ID NO: 11:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 666 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-982-785A-11

Query Match 24.7%; Score 56; DB 4; Length 666;
Best Local Similarity 43.2%; Pred. No. 2.1;
Matches 16; Conservative 8; Mismatches 11; Indels 2; Gaps 2;

QY 6 KMMVCQVGIEALVRTVLKAGDREDITEPAICALRH 42

Db 420 VSDMAVEAAELEMTROVLHAGARQDDAEPGV 451

```
RESULT 9
US-08-982-785A-8
; Sequence 8, Application US/08982785A
; Patent No. 6258929
; GENERAL INFORMATION:
; APPLICANT: Kosik, Kenneth S.
; APPLICANT: Zhou, Jianhua
; TITLE OF INVENTION: ALARM RELATED PEPTIDES AND
; NUCLEIC ACIDS AND DIAGNOSIS USING THEM
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/982,785A
; FILING DATE: 02-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/031,556
; FILING DATE: 02-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 05311/018001
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 686 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-982-785A-8

Query Match 22.2%; Score 50.5; DB 4; Length 686;
Best Local Similarity 39.5%; Pred. No. 16;
Matches 17; Conservative 6; Mismatches 19; Indels 1; Gaps 1;

QY 2 NYKNKMWCVQGGIEALVRLVLRAGDREDITEPAICALRHLS 44
| | | | | | | | | | | | | | | | | | | | | |
Db 188 NDDNKIALKNCGGIPALVRLRLKRTDLE-IRELVTGVLNLS 229

RESULT 10
US-08-982-785A-2
; Sequence 2, Application US/08982785A
; Patent No. 6258929
; GENERAL INFORMATION:
; APPLICANT: Kosik, Kenneth S.
; APPLICANT: Zhou, Jianhua
; TITLE OF INVENTION: ALARM RELATED PEPTIDES AND
; NUCLEIC ACIDS AND DIAGNOSIS USING THEM
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
```

```
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/982,785A
; FILING DATE: 02-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/031,556
; FILING DATE: 02-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 05311/018001
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 756 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-982-785A-2

Query Match 22.2%; Score 50.5; DB 4; Length 756;
Best Local Similarity 39.5%; Pred. No. 18;
Matches 17; Conservative 6; Mismatches 19; Indels 1; Gaps 1;

QY 2 NYKNKMWCVQGGIEALVRLVLRAGDREDITEPAICALRHLS 44
| | | | | | | | | | | | | | | | | | | | | |
Db 192 NDDNKIALKNCGGIPALVRLRLKRTDLE-IRELVTGVLNLS 233

RESULT 11
5284931-7 5284931
; Patent No. 5284931
; APPLICANT: SPRINGER, TIMOTHY A.; ROPHLEIN, ROBERT; MARLIN,
; STEVEN D.; DUSTIN, MICHAEL L.
; TITLE OF INVENTION: INTERCELLULAR ADHESION MOLECULES AND
; THEIR BINDING LIGANDS
; NUMBER OF SEQUENCES: 41
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/515,478
; FILING DATE: 27-APR-1990
; SEQ ID NO: 7
; LENGTH: 79
5284931-7

Query Match 21.6%; Score 49; DB 6; Length 79;
Best Local Similarity 40.0%; Pred. No. 1.7;
Matches 14; Conservative 5; Mismatches 14; Indels 2; Gaps 1;

QY 4 KKKMMVQGV--GGIEALVRLVLRAGDREDITEPAI 36
| | | | | | | | | | | | | | | | | | | | | |
Db 14 KNLTRCQVEGGAPRANLTVLLRGEKELKREPAV 48

RESULT 12
US-08-721-488-8
; Sequence 8, Application US/08721488
; Patent No. 5965388
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: Lavallie, Edward
; APPLICANT: Racie, Lisa
```


APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Spaulding, Vikki
APPLICANT: Bowman, Michael
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
TITLE OF INVENTION: ENCODING THEM
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/721,488
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 131 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-721-488-8

Query Match 21.6%; Score 49; DB 2; Length 131;
Best Local Similarity 36.4%; Pred. No. 3.3;
Matches 12; Conservative 7; Mismatches 10; Indels 4; Gaps 1;

OY 2 NYKNKMWCVQGGIEALVTVLRAGDREDITEP 34
Db 45 NYEN----IPVDKVEIETATVMPGNRELLTTP 73

RESULT 13
5489533-4
; Patent No. 5489533
; APPLICANT: SPRINGER, TIMOTHY A.; STAUNTON, DONALD E.;
; DUSTIN, MICHAEL L.
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
; ICAM-2
; NUMBER OF SEQUENCES: 16
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/384,814
; FILING DATE: 06-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 89,3075
; FILING DATE: 12-JUL-1995
; APPLICATION NUMBER: 89,307
; FILING DATE: 12-JUL-1993
; APPLICATION NUMBER: 454,294
; FILING DATE: 22-DEC-1989
; APPLICATION NUMBER: 45,963
; FILING DATE: 04-MAY-1987
; APPLICATION NUMBER: 115,798
; FILING DATE: 02-NOV-1987
; APPLICATION NUMBER: 155,943
; FILING DATE: 16-FEB-1988
; APPLICATION NUMBER: 189,815

FILING DATE: 03-MAY-1988
APPLICATION NUMBER: 250,446
FILING DATE: 28-SEP-1988
SEQ ID NO: 4:
LENGTH: 185
5489533-4

Query Match 21.6%; Score 49; DB 6; Length 185;
Best Local Similarity 40.0%; Pred. No. 5.2;
Matches 14; Conservative 5; Mismatches 14; Indels 2; Gaps 1;

OY 14 KKNKMWCVQ--GGIEALVTVLRAGDREDITEPAI 36
Db 102 KNLTRCQVEGGAPRANLTIVLLRGEKELKREPAV 136

RESULT 14
US-08-719-697-10
; Sequence: 10, Application US/08719697
; Patent No. 5928888
; GENERAL INFORMATION:
; APPLICANT: Whitney, Michael A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SENSITIVE
; TITLE OF INVENTION: AND RAPID, FUNCTIONAL IDENTIFICATION OF GENOMIC
; TITLE OF INVENTION: POLYNUCLEOTIDES AND SECONDARY SCREENING CAPABILITIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/719,697
FILING DATE: 26-SEP-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 08366/003001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 262 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-719-697-10

Query Match 21.6%; Score 49; DB 2; Length 262;
Best Local Similarity 29.4%; Pred. No. 8.1;
Matches 10; Conservative 11; Mismatches 11; Indels 2; Gaps 1;

OY 1 NNYKNKMWCVQGGIEALVTVLRAGDREDITEP 34
Db 100 DNAAQNLIKQGGPESLKKELRKIGD--EVTNP 131

RESULT 15
US-08-727-616A-10
; Sequence 10, Application US/08727616A
; Patent No. 6291162
; GENERAL INFORMATION:

APPLICANT: Tsiien, Roger Y.
APPLICANT: Zlokarnik, Gregor
TITLE OF INVENTION: SUBSTRATES FOR BETA-LACTAMASE
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/727,616A
FILING DATE: 15-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,544
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07257/034001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 262 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-727-616A-10

Query Match 21.6%; Score 49; DB 4; Length 262;
Best Local Similarity 29.4%; Pred. No. 8.1;
Matches 10; Conservative 11; Mismatches 11; Indels 2; Gaps 1;

QY 1 NNYKNMVCQVGIEALVRLRAGDREDITEP 34
Db 100 DNAQNLIKQIGGPELKKELRRIGD--EVTNP 131

Search completed: July 29, 2002, 16:05:12
Job time: 482 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 16:09:16 ; Search time 70.63 seconds
(without alignments)

59,860 Million cell updates/sec

Title: US-09-641-104a-11

Perfect score: 227

Sequence: 1 NNYKNKMVCGVGIEALVR.....AGDREDITEPAICALRHLS 44

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR71.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	227	100.0	781	2 S35099	beta-catenin - Afr
2	227	100.0	781	2 A38973	beta-catenin - hum
3	227	100.0	781	2 S35091	beta-catenin - mou
4	200	88.1	817	2 S33793	hypothetical prote
5	187	82.4	843	2 T12689	armadillo segment
6	181	79.7	806	2 JC4835	beta-catenin - Hyd
7	171	75.3	820	2 S33794	hypothetical prote
8	166	73.1	621	2 S35092	plakoglobin - mous
9	166	73.1	744	2 A32905	plakoglobin, desmo
10	152	67.0	738	2 S35093	plakoglobin - Afri
11	131	57.7	678	2 T23341	beta-catenin - Cae
12	74	32.6	811	2 T43175	armadillo protein
13	60	26.4	578	2 S50446	VAC8 protein - yea
14	59.5	26.2	825	2 T48431	hypothetical prote
15	58.5	25.8	509	2 T07721	hypothetical prote
16	57.5	25.3	381	2 AF1200	aminotransferases
17	56.5	24.9	381	2 AD1558	probable iron-cont
18	54.5	24.0	393	2 A83503	hypothetical prote
19	54.5	24.0	984	2 H90029	H+-transporting AT
20	54	23.8	1011	1 A45598	H+-transporting AT
21	53.5	23.6	609	2 S72845	probable ATPase -
22	53.5	23.6	609	2 F70512	chemotaxis respons
23	53	23.3	120	2 B72346	conserved hypothet
24	53	23.3	564	2 T37934	hypothetical prote
25	53	23.3	727	2 T47638	hypothetical prote
26	52.5	23.1	380	2 T29445	gene p120 protein
27	52.5	23.1	911	2 S28498	hypothetical prote
28	52.5	23.1	1050	2 T26395	hypothetical prote
29	52	22.9	291	2 A72711	hypothetical prote

unknown protein F2
transcription-repa
367K tegument prot
arginine decarboxy
arginine decarboxy
hypothetical prote
arginine decarboxy
chemotaxis signal
hypothetical prote
conserved hypothet
hypothetical prote
probable translati
adenylate cyclase
cyclolysin - Borde
variant surface gl
neural plakophilin

ALIGNMENTS

RESULT 1

S35099

beta-catenin - African clawed frog

C:Species: Xenopus laevis (African clawed frog)

C:Date: 10-Dec-1993.#sequence_revision 10-Nov-1995 #text_change 21-Jul-2000

C:Accession: S35099

R:McCreary, P.D.; Turck, C.W.; Gumbiner, B.

Science, 254, 1359-1361, 1991

A:Title: A homolog of the armadillo protein in Drosophila (plakoglobin) associated wi

A:Reference number: S35099; MUID:92073903

A:Accession: S35099

A:Molecule type: mRNA

A:Residues: 1-781 <MCC>

A:Cross-references: GB:M77013; NID:g214020; PIDN:AAA49670.1; PID:g214021

C:Keywords: cytoskeleton

Query Match 100.0%; Score 227; DB 2; Length 781;

Best Local Similarity 100.0%; Pred. No. 8.9e-23;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NNYKNKMVCGVGIEALVRVTLRAGDREDITEPAICALRHLS 44

Db 430 NNYKNKMVCGVGIEALVRVTLRAGDREDITEPAICALRHLS 473

RESULT 2

A38973

beta-catenin - human

C:Species: Homo sapiens (man)

C:Date: 26-Jan-1996.#sequence_revision 08-Feb-1996 #text_change 08-Oct-1999

C:Accession: A38973; S53556; S31988

R:Huelsken, J.; Birchmeier, W.; Behrens, J.

J. Cell Biol. 127, 2061-2069, 1994

A:Title: E-cadherin and APC compete for the interaction with beta-catenin and the cyt

A:Reference number: A38973; MUID:95105247

A:Accession: A38973

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-781 <HUE>

A:Cross-references: GB:Z19054; NID:g38519; PIDN:CAA79497.1; PID:g38520

A:Experimental source: placenta

R:Nollet, F.; Berx, G.; Molemans, F.; van Roy, F.

submitted to the EMBL Data Library, June 1995

A:Description: H. sapiens beta-catenin mRNA.

A:Reference number: S55356

A:Accession: S55356

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-781 <NOL>

A:Cross-references: EMBL:X87838; NID:g1154853; PIDN:CAA61107.1; PID:g860988

C:Comment: The cytoplasmic protein beta-catenin normally binds to a region in the cyt

microfilament network.

C:Comment: Cellular levels of beta-catenin are regulated in part by the adenomatous polyd cancerous cell growth.

C:Genetics:

A:Gene: GDB:CTNNB1; CTNNB

A:Cross-references: GDB:I41922; OMIM:116806

A:Map position: 3p22-3p21.3

C:Keywords: apoptosis; carcinogenesis; cell adhesion; cytosol

F:151-676/Region: 40-residue repeats

Query Match 100.0%; Score 227; DB 2; Length 781;
Best Local Similarity 100.0%; Pred. No. 8.9e-23;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNYKNKMMVCGVGIEALVTVLRAGDREDITEPAICALRHLS 44

DB 430 NNYKNKMMVCGVGIEALVTVLRAGDREDITEPAICALRHLS 473

RESULT 3

S35091

beta-catenin - mouse

C:Species: Mus musculus (house mouse)

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999

C:Accession: S35091

R:Butz, S.; Stappert, J.; Weissig, H.; Kemler, R.

A:Title: Plakoglobin and beta-catenin: distinct but closely related.

A:Reference number: S35091; MUID:92376536

A:Accession: S35091

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-781 <BUT>

A:Cross-references: EMBL:M90364; NID:g192141; PIDN:AAA37280.1; PID:g192142

C:Keywords: cytoskeleton

Query Match 100.0%; Score 227; DB 2; Length 781;
Best Local Similarity 100.0%; Pred. No. 8.9e-23;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNYKNKMMVCGVGIEALVTVLRAGDREDITEPAICALRHLS 44

DB 430 NNYKNKMMVCGVGIEALVTVLRAGDREDITEPAICALRHLS 473

RESULT 4

S33793

hypothetical protein - spoonworm (Urechis caupo)

C:Species: Urechis caupo

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Aug-1997

C:Accession: S33793

R:Rosenthal, E.

A:Title: Biophys. Acta 1173, 337-341, 1993

A:Title: Identification of homologues to beta-catenin/plakoglobin/armadillo in two inver

A:Reference number: S33793; MUID:93305730

A:Accession: S33793

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-817 <ROS>

Query Match 88.1%; Score 200; DB 2; Length 817;
Best Local Similarity 84.1%; Pred. No. 4.5e-19;
Matches 37; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 NNYKNKMMVCGVGIEALVTVLRAGDREDITEPAICALRHLS 44

DB 452 NNORKKTVTCQVGIEALVTVLRAGDREDITEPAICALRHLS 495

RESULT 5

S33794

hypothetical protein - sea urchin (Tripneustes gratilla)

T12689

armadillo segment polarity protein - fruit fly (Drosophila melanogaster)

N:Alternate names: protein 86E4.6

C:Species: Drosophila melanogaster

C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 21-Jul-2000

C:Accession: T12689; A31861

R:Ferraz, C.; Vidal, S.; Brun, C.; Bucheton, A.; Demaille, J.G.

submitted to the EMBL Data Library, January 1998

A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.

A:Reference number: Z17572

A:Accession: T12689

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-843 <FER>

A:Cross-references: EMBL:AL021106; NID:el371406; PID:el249776; PIDN:CAA15946.1

A:Experimental source: clone cosmid 63B12

R:Rigleman, B.; Wieschaus, E.; Schedl, P.

Genes Dev. 3, 96-113, 1989

A:Title: Molecular analysis of the Armadillo locus: uniformly distributed transcripts

A:Reference number: A31861; MUID:89211895

A:Accession: A31861

A:Molecule type: DNA

A:Residues: 1-843 <RIG>

A:Cross-references: EMBL:X54468; NID:g7610; PIDN:CAA38350.1; PID:g7611

C:Genetics:

A:Gene: arm

A:Cross-references: FlyBase:FBgn0000117

A:Introns: 11/1; 557/3; 624/3; 715/3; 748/1

A>Note: 86E4.6

C:Keywords: cytoskeleton

Query Match 82.4%; Score 187; DB 2; Length 843;
Best Local Similarity 75.0%; Pred. No. 2.8e-17;
Matches 33; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 1 NNYKNKMMVCGVGIEALVTVLRAGDREDITEPAICALRHLS 44

DB 438 NNORKKATVCQVGVDALVTVLRAGDREDITEPAICALRHLS 481

RESULT 6

JC4835

beta-catenin - Hydra magnipapillata

C:Species: Hydra magnipapillata

C:Date: 15-Aug-1996 #sequence_revision 15-Oct-1996 #text_change 21-Jul-2000

C:Accession: JC4835

R:Hobmayer, E.; Hatta, M.; Fischer, R.; Fujisawa, T.; Holstein, T.W.; Sugiyama, T.

Gene 172, 155-159, 1996

A:Title: Identification of a Hydra homologue of the beta-catenin/plakoglobin/armadill

A:Reference number: JC4835; MUID:96257271

A:Accession: JC4835

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-806 <HOB>

A:Cross-references: GB:U36781; NID:g1407600; PIDN:AAC47137.1; PID:g1407601

C:Comment: This protein plays roles in cadherin-mediated cell adhesion and in signal

C:Genetics:

A:Gene: betaCtn

Query Match 79.7%; Score 181; DB 2; Length 806;
Best Local Similarity 77.3%; Pred. No. 1.8e-16;
Matches 34; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 NNYKNKMMVCGVGIEALVTVLRAGDREDITEPAICALRHLS 44

DB 489 NNPRNKQVVQVGIEALVTVLRAGDREDITEPAICALRHLS 532

RESULT 7

S33794

hypothetical protein - sea urchin (Tripneustes gratilla)


```
Best Local Similarity: 25.6%; Pred.No. 4; 1;  
Matches: 11; Conservative 16; Mismatches 14; Indels 2; Gaps 1;
```

QY 12 N Y K N K M V Q V G T E A L V R T V L R A G D R I T E P A I C A L R H L T S 44
DB : | | | | : | | : | | : | : | :
 115 N N E N K L L I V M G G L E P L I N Q M M - G D N V E O C N A V G C I T N L A T 155

RESULT #14
T48431.....
hypothetical protein F8F6_250 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 20-Apr-2000;#sequence_revision 20-Apr-2000 #text_change 28-Jul-2000
C:Accession: T48431.
R:Bévan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft
submitted to The Protein Sequence Database, March 2000
A:Reference number: Z24488
A:Accession: T48431.
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-825 <BEV>
A:Cross-references: EMBL:AL162873
A:Experimental source: cultivar Columbia; BAC clone F8F6
C:Genetics:
A:Map position: 5
A:Introps.: 474/3; 522/3
A>Note: F8F6_250
C:Superfamily: Arabidopsis thaliana hypothetical protein F8F6_250

Query Match 26.2% ; Score 59.5; DB 2; Length 825;
Best Local Similarity . . 41.2% ; Pred. No. 6.3;
Matches: 14; Conservative 8; Mismatches 11; Indels 1; Gaps 1;

QY 11 Q V G G I E A L V R T V L R A G D R I T E P A I C A L R H L T S 44
DB : | | | : | : | | | : | | : | | | |
 298 Q L G V F S I V K R V M T Q G A L H D I R Q - L Q C M L R N L T S 330

```

RESULT 15
T07721
hypothetical protein T23J7.150 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 22-Oct-1999
C:Accession: T07721
R:Quetier, F.; Choisme, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.
submitted to the Protein Sequence Database, April 1999
A:Reference number: 215793
A:Accession: T07721
A:Molecule type: DNA
A:Residues: 1-509 <QUE>
A:Cross-references: EMBL:AL049746; GSPDB:GN00061; ATSP:T23J7.150
A:Experimental source: cultivar Columbia; BAC clone T23J7
C:Genetics:
A:Gene: ATSP:T23J7.150
A:Map position: 3

Query Match 25.8% Score 58.5; DB 2; Length 509;
Best Local Similarity 35.08; Pred. No. 5.8;
Matches 14; Conservative 10; Mismatches 15; Indels 1; Gaps 1;

Qy 4 KKKMMVCQVGIGIEALVTVLRAGDREDITEPAICALRHLT 43
Db 298 ENKMWIGVLGAVEPLIH-ALRSSSESRARQDAALALYHLS 336

Search completed: July 29, 2002, 16:09:17
Job time: 727 sec

```

Search completed: July 29, 2002, 16:09:17
Job time: 727 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 16:09:55 ; Search time 34.24 Seconds
(without alignments)
49.756 Million cell updates/sec

Title: US-09-641-104A-11

Perfect score: 227

Sequence: 1 NNYKNMVMQVGGIEALVR.....AGDREDITEPAICALRLHLS 44

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	227	100.0	781	1 CTNB_HUMAN	P35222 homo sapien
2	227	100.0	781	1 CTNB_MOUSE	Q02248 mus musculu
3	227	100.0	781	1 CTNB_RAT	Q9wu82 rattus norv
4	227	100.0	781	1 CTNB_XENLA	P26233 xenopus lae
5	200	88.1	818	1 CTNB_DRECA	P35224 urechis cau
6	187	82.4	843	1 ARM_DROME	P18824 drosophila
7	184	81.1	813	1 ARM_MUSDO	Q02453 musca domes
8	171	75.3	820	1 CTNB_TRIGR	P02453 tripneustes
9	166	73.1	621	1 PLAK_MOUSE	Q02257 mus musculu
10	166	73.1	743	1 PLAK_HUMAN	P14923 homo sapien
11	152	67.0	738	1 PLAK_XENLA	P30998 xenopus lae
12	60	26.4	578	1 VAC8_YEAST	P39968 saccharomyc
13	58.5	25.8	797	1 PRP3_HUMAN	Q9y446 homo sapien
14	54	23.8	1011	1 ATC_TRYBB	P35315 trypanosoma
15	53.5	23.6	609	1 YL15_MYCLE	P46509 mycobacteri
16	53.5	23.6	609	1 YL15_MYCTU	Q33250 mycobacteri
17	53	23.3	120	1 CHEY_THEMA	Q56312 thermotoga
18	52.5	23.1	911	1 CTDI_MOUSE	P30999 mus musculu
19	52.5	23.1	968	1 CTDI_HUMAN	Q60716 homo sapien
20	52	22.9	291	1 PNK_AERPE	Q9y408 aeopryum p
21	52	22.9	3421	1 TEGU_HSVB	P28955 equine herp
22	51.5	22.7	797	1 PRP3_MOUSE	Q9y423 mus musculu
23	51.5	22.7	962	1 ARVC_HUMAN	Q00192 homo sapien
24	51	22.5	683	1 EFG2_TREPA	Q83464 treponema p
25	51	22.5	1211	1 PRP4_HUMAN	Q99569 homo sapien
26	51	22.5	1705	1 CYAA_BORBR	Q57506 bordetella
27	51	22.5	1706	1 CYAA_BORPE	P15318 bordetella
28	50	22.0	302	1 GPRR_AGRTU	P25547 agrobacteri
29	50	22.0	616	1 PRPA_SOLIU	P21342 solanum tub
30	49.5	21.8	228	1 R55_DROME	Q24186 drosophila
31	49.5	21.8	325	1 ESKI_TRYBB	P04477 trypanosoma
32	49	21.6	307	1 BLAC_BACLI	P00808 bacillus li
33	49	21.6	449	1 AK3_ECOLI	P08660 escherichia

RESULT

ID	CTNB_HUMAN	STANDARD;	PRT;	781 AA.
AC	P35222			
DT	01-FEB-1994 (Rel. 28, Created)			
DT	01-FEB-1994 (Rel. 28, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Beta-catenin.			
GN	CTNNB1 OR CTNNB			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Placenta;			
RC	MEDLINE=95105247; PubMed=7806582;			
RA	Huelsken J., Birchmeier W., Behrens J.;			
RT	"E-cadherin and APC compete for the interaction with beta-catenin and the cytoskeleton."			
RL	J. Cell Biol. 127:2061-2069(1994).			
RN	[2]			
RP	REVIEW.			
RX	MEDLINE=20145417; PubMed=10679188;			
RA	Kikuchi A.;			
RT	"Regulation of beta-catenin signaling in the Wnt pathway."			
RL	Biochem. Biophys. Res. Commun. 268:243-248(2000).			
RN	[3]			
RP	X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 133-664.			
RX	MEDLINE=20578006; PubMed=11136974;			
RA	Graham T.A., Weaver C., Mao F., Kimelman D., Xu W.;			
RT	"Crystal structure of a beta-catenin/Tcf complex."			
RL	Cell 103:885-896(2000).			
CC	-1- FUNCTION: INVOLVED IN THE REGULATION OF CELL ADHESION AND IN SIGNAL TRANSDUCTION THROUGH THE WNT PATHWAY.			
CC	-1- SUBUNIT: THE CYTOSOLIC FORM BOUND TO AXIN ALLOWS THE STIMULATION OF ITS DEGRADATION. THE CYTOSOLIC FORM ALSO INTERACTS WITH APC AND "E-CADHERIN. THE NUCLEAR FORM BINDS TCF/LEF-1 AND MAY ALSO BIND PONTIN2 AND DUPLIN. CAN HETERODIMERIZE WITH ALPHA-CATENIN AND ALSO FORM A TRIMERIC COMPLEX WITH CADHERIN. MEDIATES THE ANCHORAGE OF THE CADHERIN COMPLEX TO THE CORTICAL ACTIN CYTOSKELETON.			
CC	-1- SUBCELLULAR LOCATION: CYTOPLASMIC WHEN IT IS UNSTABILIZED (HIGH LEVEL OF PHOSPHORYLATION). TRANSLOCATES TO THE NUCLEUS WHEN IT IS STABILIZED (LOW LEVEL OF PHOSPHORYLATION).			
CC	-1- PTM: PROBABLY PHOSPHORYLATED BY GSK-3B. WNT-DEPENDENT ACTIVATION OF DVL ANTAGONIZES THE ACTION OF GSK-3B. PHOSPHORYLATION OF BETA-CATENIN IS REDUCED AND THE PROTEIN IS NO LONGER DEGRADED, RESULTING IN ITS ACCUMULATION IN CYTOPLASM.			
CC	-1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.			
CC	-1- SIMILARITY: CONTAINS 12 ARM REPEATS.			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use, by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

P05362 homo sapien
O58974 pyrococcus
Q9kbe4 bacillus ba
P29139 paenibacill
P18667 synecococc
Q96412 dianthus ca
P98260 mus musculu
Q926q0 chlamydia p
Q58588 methanococc
Q92na9 helicobacte
O25997 helicobacte
Q00127 ictalurid h

entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>) or send an email to license@isb-sib.ch.

```
CC CC EMBL; X87838; CAA61107.1; -.
CC CC EMBL; Z19054; CAA79497.1; -.
CC CC PIR; S31988; S31988.
CC CC PDB; 1G3J; 17-JAN-01.
CC CC TRANSFAC; T02872; -.
CC CC MIN; 116806; -.
CC CC InterPro; IPR000225; Armadillo.
CC CC Pfam; PF00514; Armadillo_seg; 12.
CC CC SMART; SM00185; ARM; 11.
CC CC PROSITE; PS50176; ARM_REPEAT; 9.
CC CC Cell adhesion; Cytoskeleton; Structural protein; Nuclear protein;
KW Repeat; 3D-structure.
FT REPEAT 151 191 ARM 1.
FT REPEAT 193 234 ARM 2.
FT REPEAT 235 276 ARM 3.
FT REPEAT 277 318 ARM 4.
FT REPEAT 319 360 ARM 5.
FT REPEAT 361 389 ARM 6.
FT REPEAT 400 441 ARM 7.
FT REPEAT 442 484 ARM 8.
FT REPEAT 489 530 ARM 9.
FT REPEAT 531 571 ARM 10.
FT REPEAT 594 636 ARM 11.
FT REPEAT 637 666 ARM 12.
SQ SEQUENCE 781 AA; 85496 MW; CB78F165A3BEF68E CRC64;

Query Match 100.0%; Score 227; DB 1; Length 781;
Best Local Similarity 100.0%; Pred. No. 1.9e-23;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NNYKNKMMVCGVGIEALVRLVLRAGREDITEPAICALRHLS 44
Db 430 NNYKNKMMVCGVGIEALVRLVLRAGREDITEPAICALRHLS 473

RESULT 2
CTNB_MOUSE STANDARD; PRT; 781 AA.
AC Q02248;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Beta-catenin.
GN CTNNB1 OR CATNB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92376536; PubMed=1509266;
RA Butz S., Stappert J., Weissig H., Kemler R.;
RT "Plakoglobin and beta-catenin: distinct but closely related.";
RL Science 257:1142-1144(1992).
RN [2]
RP REVIEW.
RX MEDLINE=20145417; PubMed=10679188;
RA Kikuchi A.;
RT "Regulation of beta-catenin signaling in the Wnt pathway.";
RL Biochem. Biophys. Res. Commun. 268:243-248(2000).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 150-665.
RX MEDLINE=97442350; PubMed=9298899;
RA Huber A.H., Nelson W.J., Weis W.I.;
RT "Three-dimensional structure of the armadillo repeat region of beta-catenin.";
RL Cell 90:871-882(1997).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 118-149 IN COMPLEX WITH
```

```
RP CTNNB1.
RX MEDLINE=20337986; PubMed=10882138;
RA Pokutta S., Weis W.I.;
RT "Structure of the dimerization and beta-catenin-binding region of
alpha-catenin.";
RL Mol. Cell 5:533-543(2000).
CC -1- FUNCTION: INVOLVED IN THE REGULATION OF CELL ADHESION AND IN
CC SIGNAL TRANSDUCTION THROUGH THE WNT PATHWAY.
CC -1- SUBUNIT: THE CYTOSOLIC FORM BOUND TO AXIN ALLOWS THE STIMULATION
CC OF ITS DEGRADATION. THE CYTOSOLIC FORM ALSO INTERACTS WITH APC AND
CC E-CADHERIN. THE NUCLEAR FORM BINDS TCF/LEF-1 AND MAY ALSO BIND
CC PONTIN52 AND DUPLIN. CAN HETERODIMERIZE WITH ALPHA-CATENIN AND
CC ALSO FORM A TRIMERIC COMPLEX WITH CADHERIN. MEDIATES THE ANCHORAGE
CC OF THE CADHERIN COMPLEX TO THE CORTICAL ACTIN CYTOSKELETON.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC WHEN IT IS UNSTABILIZED (HIGH
CC LEVEL OF PHOSPHORYLATION). TRANSLOCATES TO THE NUCLEUS WHEN IT IS
CC STABILIZED (LOW LEVEL OF PHOSPHORYLATION).
CC -1- PTM: PROBABLY PHOSPHORYLATED BY GSK-3B. WNT-DEPENDENT ACTIVATION
CC OF DVL ANTAGONIZES THE ACTION OF GSK-3B. PHOSPHORYLATION OF BETA-
CC CATENIN IS REDUCED AND THE PROTEIN IS NO LONGER DEGRADED,
CC RESULTING IN ITS ACCUMULATION IN CYTOPLASM.
CC -1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -1- SIMILARITY: CONTAINS 12 ARM REPEATS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC EMBL; M90364; AAA37280.1; -.
CC PIR; S35091; S35091.
CC PDB; 2BCT; 15-OCT-97.
CC PDB; 3BCT; 19-NOV-97.
CC PDB; 1DOW; 12-JUL-00.
CC TRANSFAC; T02984; -.
CC MGD; MGI:88276; Catnb.
CC InterPro; IPR000225; Armadillo.
CC Pfam; PF00514; Armadillo_seg; 12.
CC SMART; SM00185; ARM; 11.
CC PROSITE; PS50176; ARM_REPEAT; 9.
CC Cell adhesion; Cytoskeleton; Structural protein; Nuclear protein;
KW Repeat; 3D-structure.
FT REPEAT 151 191 ARM 1.
FT REPEAT 193 234 ARM 2.
FT REPEAT 235 276 ARM 3.
FT REPEAT 277 318 ARM 4.
FT REPEAT 319 360 ARM 5.
FT REPEAT 361 389 ARM 6.
FT REPEAT 400 441 ARM 7.
FT REPEAT 442 484 ARM 8.
FT REPEAT 489 530 ARM 9.
FT REPEAT 531 571 ARM 10.
FT REPEAT 594 636 ARM 11.
FT REPEAT 637 666 ARM 12.
SQ SEQUENCE 781 AA; 85470 MW; D708F170A3FBED6E CRC64;

Query Match 100.0%; Score 227; DB 1; Length 781;
Best Local Similarity 100.0%; Pred. No. 1.9e-23;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NNYKNKMMVCGVGIEALVRLVLRAGREDITEPAICALRHLS 44
Db 430 NNYKNKMMVCGVGIEALVRLVLRAGREDITEPAICALRHLS 473

RESULT 3
CTNB_MOUSE STANDARD; PRT; 781 AA.
AC Q9WUB2;
```

DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Beta-catenin.
GN CTNNB1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE-Testis;
RA MEDLINE=99428593; PubMed=10497305;
RA Chung S.S.W., Lee W.M., Cheng C.Y.;
RT "Study on the formation of specialized inter-Sertoli cell junctions in
RT vitro";
RL J. Cell. Physiol. 181:258-272(1999).
CC -!- FUNCTION: INVOLVED IN THE REGULATION OF CELL ADHESION AND IN
CC SIGNAL TRANSDUCTION THROUGH THE WNT PATHWAY (BY SIMILARITY).
CC -!- SUBUNIT: THE CYTOSOLIC FORM BOUND TO AXIN ALLOWS THE STIMULATION
CC OF ITS DEGRADATION. THE CYTOSOLIC FORM ALSO INTERACTS WITH APC AND
CC E-CADHERIN. THE NUCLEAR FORM BINDS TCF/LEF-1 AND MAY ALSO BIND
CC PONTIN52 AND DUPLIN. CAN HETERODIMERIZE WITH ALPHA-CATENIN AND
CC ALSO FORM A TRIMERIC COMPLEX WITH CADHERIN. MEDIATES THE ANCHORAGE
CC OF THE CADHERIN COMPLEX TO THE CORTICAL ACTIN CYTOSKELETON (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC WHEN IT IS UNSTABILIZED (HIGH
CC LEVEL OF PHOSPHORYLATION). TRANSLOCATES TO THE NUCLEUS WHEN IT IS
CC STABILIZED (LOW LEVEL OF PHOSPHORYLATION) (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE TESTIS.
CC -!- DEVELOPMENTAL STAGE: HIGHLY EXPRESSED AT E30-60 DAY DPC IN THE
CC TESTIS. REDUCED EXPRESSION AT E90 DAY DPC.
CC -!- PTM: PROBABLY PHOSPHORYLATED BY GSK-3B. WNT-DEPENDENT ACTIVATION
CC OF DVL ANTAGONIZES THE ACTION OF GSK-3B. PHOSPHORYLATION OF BETA-
CC CATENIN IS REDUCED AND THE PROTEIN IS NO LONGER DEGRADED.
CC RESULTING IN ITS ACCUMULATION IN CYTOPLASM (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -!- SIMILARITY: CONTAINS 12 ARM REPEATS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
DR EMBL; AF121265; AAD28504.1; -
DR HSSP; Q02248; 2BCT.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 12.
DR SMART; SM00185; ARM; 11.
DR PROSITE; PS50176; ARM_REPEAT; 9.
KW Cell adhesion; Cytoskeleton; Structural protein; Nuclear protein;
KW Repeat.
FT REPEAT 151 191 ARM 1.
FT REPEAT 193 234 ARM 2.
FT REPEAT 235 276 ARM 3.
FT REPEAT 277 318 ARM 4.
FT REPEAT 319 360 ARM 5.
FT REPEAT 361 389 ARM 6.
FT REPEAT 400 441 ARM 7.
FT REPEAT 442 484 ARM 8.
FT REPEAT 489 530 ARM 9.
FT REPEAT 531 571 ARM 10.
FT REPEAT 594 636 ARM 11.
FT REPEAT 637 666 ARM 12.
SQ SEQUENCE 781 AA; 85454 MW; 9C29186B6DD54B87 CRC64;

Query Match 100.0%; Score 227; DB 1; Length 781;
Best Local Similarity 100.0%; Pred. No. 1.9e-23;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNYKNKMMVCQVGGIEALVRLTVLRAGDRDITEPAICALRHLS 44
DB 430 NNYKNKMMVCQVGGIEALVRLTVLRAGDRDITEPAICALRHLS 473
RESULT 14
CTNB_XENLA STANDARD; PRT; 781 AA.
ID CTNB_XENLA
AC P26233;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Beta-catenin.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Piploidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RN SEQUENCE FROM N.A.; AND PARTIAL SEQUENCE.
RC TISSUE-Brain;
RX MEDLINE=92073903; PubMed=1962194;
RA McCrea P.D., Turk C.W., Gumbiner B.M.;
RT "A homolog of the armadillo protein in Drosophila (plakoglobin)
RT associated with E-cadherin";
RL Science 254:1359-1361(1991).
CC -!- FUNCTION: BINDS TO THE CYTOPLASMIC DOMAIN OF THE CELL-CELL
CC ADHESION MOLECULE E-CADHERIN, AND PERHAPS TO OTHER (MEMBRANE)
CC PROTEINS. THE ASSOCIATION OF CATENINS TO CADHERINS PRODUCES A
CC COMPLEX WHICH IS LINKED TO THE ACTIN FILAMENT NETWORK, AND WHICH
CC SEEMS TO BE OF PRIMARY IMPORTANCE FOR CADHERINS CELL-ADHESION
CC PROPERTIES. ESSENTIAL COMPONENT OF THE VERTEBRATE WNT-SIGNALING
CC PATHWAY AND IS REQUIRED FOR FORMATION OF THE EMBRYONIC AXIS.
CC -!- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -!- SIMILARITY: CONTAINS 10 ARM REPEATS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
DR EMBL; M77013; AAA49670.1; -
DR PIR; S35099; S35099.
DR HSSP; Q02248; 2BCT.
DR TRANSFAC; T03026; -
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 12.
DR SMART; SM00185; ARM; 11.
DR PROSITE; PS50176; ARM_REPEAT; 9.
KW Cell adhesion; Cytoskeleton; Structural protein; Repeat.
FT REPEAT 141 180 ARM 1.
FT REPEAT 181 225 ARM 2.
FT REPEAT 226 264 ARM 3.
FT REPEAT 267 306 ARM 4.
FT REPEAT 351 390 ARM 5.
FT REPEAT 391 429 ARM 6.
FT REPEAT 432 473 ARM 7.
FT REPEAT 479 519 ARM 8.
FT REPEAT 521 562 ARM 9.
FT REPEAT 584 623 ARM 10.
FT REPEAT 625 664 ARM 11.
SQ SEQUENCE 781 AA; 85449 MW; 3ECD27232239F799 CRC64;

Query Match 100.0%; Score 227; DB 1; Length 781;
Best Local Similarity 100.0%; Pred. No. 1.9e-23;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NNYKNKMMVCQVGGIEALVRLTVLRAGDRDITEPAICALRHLS 44
DB 430 NNYKNKMMVCQVGGIEALVRLTVLRAGDRDITEPAICALRHLS 473


```
DR PROSITE: PS50176; ARM_REPEAT; 9;
KW Developmental protein; Segmentation polarity protein; Repeat;
KW Cell adhesion; Cytoskeleton; Structural protein.
FT DOMAIN 1 148 ASP/GLU-RICH (ACIDIC).
FT REPEAT 139 178 ARM 1.
FT REPEAT 180 221 ARM 2.
FT REPEAT 223 262 ARM 3.
FT REPEAT 265 304 ARM 4.
FT REPEAT 349 387 ARM 5.
FT REPEAT 388 427 ARM 6.
FT REPEAT 430 471 ARM 7.
FT REPEAT 477 517 ARM 8.
FT REPEAT 587 626 ARM 9.
FT REPEAT 628 667 ARM 10.
SQ SEQUENCE 813 AA; 88237 MW; 73561F59BDBFA580 CRC64;

Query Match 81.1%; Score 184; DB 1; Length 813;
Best Local Similarity 72.7%; Pred. No. 1.6e-17;
Matches 32; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 1 NNYKNKMMVCQVGGIEALVRLVLRAGDREDITEPAICALRHLS 44
II :II :IIIIII:IIIIII:IIIIII:IIIIII:IIIIII:
DB 428 NNQRKATVCQVGGVDALVRTIINAGDREDITEPAICALRHLS 471

RESULT 8
CTNB_TRIGR STANDARD; PRT; 820 AA.
AC P35223;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Beta-catenin.
OS Tripeustes gratilla (Hawaiian sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Temnopleuroidea; Toxopneustidae;
OC Tripeustes.
OC NCBI_taxid=7673;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93305730; PubMed=8318544;
RA Rosenthal E.T.;
RT Identification of homologues to beta-catenin/plakoglobin/armadillo
in two invertebrates, Urechis caupo and Tripeustes gratilla.";
RL Biochim. Biophys. Acta 1173:337-341(1993).
CC -1- FUNCTION: BINDS TO THE CYTOPLASMIC DOMAIN OF THE CELL-CELL
ADHESION MOLECULE E-CADHERIN, AND PERHAPS TO OTHER (MEMBRANE)
PROTEINS. THE ASSOCIATION OF CATENINS TO CADHERINS PRODUCES A
COMPLEX WHICH IS LINKED TO THE ACTIN FILAMENT NETWORK, AND WHICH
SEEMS TO BE OF PRIMARY IMPORTANCE FOR CADHERINS CELL-ADHESION
PROPERTIES.
CC -1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -1- SIMILARITY: CONTAINS 10 ARM REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L10354; AAA30089.1;
CC PIR: S33794; S33794.
CC HSSP: Q02248; IDOW.
CC InterPro: IPR000225; Armadillo.
CC Pfam: PF00514; Armadillo_seg; 12.
CC SMART: SM00185; ARM; 12.
CC PROSITE: PS50176; ARM_REPEAT; 9.
KW Cell adhesion; Cytoskeleton; Structural protein; Repeat.
FT REPEAT 157 196 ARM 1.
FT REPEAT 199 239 ARM 2.
```

```
FT REPEAT 241 280 ARM 3.
FT REPEAT 283 322 ARM 4.
FT REPEAT 367 405 ARM 5.
FT REPEAT 406 445 ARM 6.
FT REPEAT 448 489 ARM 7.
FT REPEAT 495 535 ARM 8.
FT REPEAT 603 642 ARM 9.
FT REPEAT 644 683 ARM 10.
SQ SEQUENCE 820 AA; 89361 MW; 57255E0F57795FD3 CRC64;

Query Match 75.3%; Score 171; DB 1; Length 820;
Best Local Similarity 68.2%; Pred. No. 1e-15;
Matches 30; Conservative 11; Mismatches 3; Indels 0; Gaps 0;

QY 1 NNYKNKMMVCQVGGIEALVRLVLRAGDREDITEPAICALRHLS 44
II :II :IIIIII:IIIIII:IIIIII:IIIIII:IIIIII:
DB 446 NNSRNKMMVMSQAGVEALVQTLKAGDREDITEPAICALRHVS 489

RESULT 9
PLAK_MOUSE STANDARD; PRT; 621 AA.
AC Q02257;
DT 01-JUL-1993 (Rel. 26; Created)
DT 01-OCT-1996 (Rel. 34; Last sequence update)
DT 01-MAR-2002 (Rel. 41; Last annotation update)
DE Junction plakoglobin (Desmoplakin III) (Fragment).
GN JUP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OC NCBI_taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93376536; PubMed=1509266;
RA Butz S., Stappert J., Weissig H., Kemler R.;
RT "Plakoglobin and beta-catenin: distinct but closely related.";
RL Science 257:1142-1144(1992).
RN [2]
RP REVISIONS TO 294 AND 296.
RA Butz S.;
RA Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: COMMON JUNCTIONAL PLAQUE PROTEIN. THE MEMBRANE-
ASSOCIATED PLAQUES ARE ARCHITECTURAL ELEMENTS IN AN IMPORTANT
STRATEGIC POSITION TO INFLUENCE THE ARRANGEMENT AND FUNCTION
OF BOTH THE CYTOSKELETON AND THE CELLS WITHIN THE TISSUE. THE
PRESENCE OF PLAQUEGLIBIN IN BOTH THE DESMOSOMES AND IN THE
INTERMEDIATE JUNCTIONS SUGGESTS THAT IT PLAYS A CENTRAL ROLE IN
THE STRUCTURE AND FUNCTION OF SUBMEMBRANOUS PLAQUES.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC IN A SOLUBLE AND MEMBRANE-
ASSOCIATED FORM...
CC -1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -1- SIMILARITY: CONTAINS AT LEAST 9 ARM REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: W90365; AAB02885.1;
CC PIR: S35092; S35092.
CC HSSP: Q02248; IDOW.
CC MGD: MGI:96650; Jup.
CC InterPro: IPR00225; Armadillo.
CC Pfam: PF00514; Armadillo_seg; 11.
CC SMART: SM00185; ARM; 8.
CC PROSITE: PS50176; ARM_REPEAT; 9.
KW Cell adhesion; Cytoskeleton; Structural protein; Repeat.
```

.....

[illegible]

or send

DR EMBL; M95593; AAA49931.1; -;
DR EMBL; X67078; CAA47463.1; -;
DR PIR; S24636; S24636.
DR PIR; S35093; S35093.
DR HSP; Q02248; 2BCT.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 11.
DR SMART; SM00185; ARM; 10.
DR PROSITE; PS01176; ARM_REPEAT; 9.
KW Cell adhesion; Cytoskeleton; Structural protein; Repeat.
FT REPEAT 128 167 ARM 1.
FT REPEAT 212 251 ARM 2.
FT REPEAT 251 293 ARM 3.
FT REPEAT 338 377 ARM 4.
FT REPEAT 379 416 ARM 5.
FT REPEAT 419 460 ARM 6.
FT REPEAT 466 506 ARM 7.
FT REPEAT 508 547 ARM 8.
FT REPEAT 570 609 ARM 9.
FT REPEAT 611 650 ARM 10.
FT CONFLICT 185 185 I -> V (IN REF. 2).
FT CONFLICT 226 226 A -> T (IN REF. 2).
SQ SEQUENCE 738 AA; 81711 MW; 569DBE69D08B58 CRC64;

Query Match 67.0%; Score 152; DB 1; Length 738;
Best Local Similarity 59.1%; Pred. No. 3.7e-13;
Matches 26; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

QY 1 NYKNKMYCVQGVGTEALVTVLRAGDREDITEPAICALRHLS 44
DB 417 NGRNKTTLVTSQVGSLEHTILRASDKDDIAEPVAVCAMRHLS 460

RESULT 12
VAC8_YEAST
ID VAC8_YEAST STANDARD; PRT; 578 AA.
AC P39968;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Vacuolar protein 8.
GN VAC8 OR YEL013W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / AB972;
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
RA Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,
RA Hymen R., Kayser A., Komp C., Laskari D., Lew H., Lin D.,
RA Mosedale D., Nakamura K., Namach A., Norgren R., Oefner P., Oh C.,
RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP CHARACTERIZATION.
RA Wang Y.X., Catlett N.L., Weisman L.S.;
RL Unpublished observations (XXX-1997).
CC -!- FUNCTION: FUNCTIONS IN BOTH VACUOLE INHERITANCE AND PROTEIN
CC TARGETING FROM THE CYTOPLASM TO VACUOLE.
CC -!- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY. MOST SIMILAR TO
CC PLAKOGLOBINS
CC -!- SIMILARITY: CONTAINS 9 ARM REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U18530; AAB64490.1; -;
DR SGD; S0000739; VAC8.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 9.
DR SMART; SM00185; ARM; 8.
DR PROSITE; PS01176; ARM_REPEAT; 7.
KW Cytoskeleton; Repeat.
FT REPEAT 37 75 ARM 1.
FT REPEAT 76 114 ARM 2.
FT REPEAT 116 155 ARM 3.
FT REPEAT 157 196 ARM 4.
FT REPEAT 198 237 ARM 5.
FT REPEAT 239 280 ARM 6.
FT REPEAT 282 321 ARM 7.
FT REPEAT 323 363 ARM 8.
FT REPEAT 407 446 ARM 9.
SQ SEQUENCE 578 AA; 63207 MW; B2E75774B47E5933 CRC64;

Query Match 26.4%; Score 60; DB 1; Length 578;
Best Local Similarity 25.6%; Pred. No. 1.3;
Matches 11; Conservative 16; Mismatches 14; Indels 2; Gaps 1;

QY 2 NYKNKMYCVQGVGTEALVTVLRAGDREDITEPAICALRHLS 44
DB 115 NNENKLLIVEMGSLIPLNQMM--GDNVEVQCNAVGCITNLAT 155

RESULT 13
PKP3_HUMAN
ID PKP3_HUMAN STANDARD; PRT; 797 AA.
AC Q9Y446;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Plakophilin 3.
GN PKP3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Colon carcinoma;
RX MEDLINE-99302495; PubMed-10374265;
RA Schmidt A., Langbein L., Praetzel S., Rode M., Rackwitz H.-R.,
RA Franke W.W.;
RT "Plakophilin 3 - a novel cell-type-specific desmosomal plaque
RT protein."
RL Differentiation 64:291-306(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-99310799; PubMed-10381383;
RA Bonne S., van Hengel J., Nollet F., Kools P., van Roy P.;
RT "Plakophilin-3, a novel armadillo-like protein present in nuclei and
RT desmosomes of epithelial cells."
RL J. Cell Sci. 112:2265-2276(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE-Placenta;
RA Strausberg R.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MAY PLAY A ROLE IN JUNCTIONAL PLAQUES.
CC -!- SUBCELLULAR LOCATION: Nuclear and associated with desmosomes.
CC -!- TISSUE SPECIFICITY: FOUND IN DESMOSOMES OF MOST SIMPLE AND
CC STRATIFIED EPITHELIA. NOT FOUND IN FORESKIN FIBROBLASTS AND
CC VARIOUS SARCOMA-DERIVED CELL LINES. BESIDE DENSITIC RETICULAR
CC CELLS OF LYMPHATIC FOLLICLES NOT FOUND IN NON-EPITHELIAL
CC DESMOSOME-BEARING TISSUES.
CC -!- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 16:12:09 ; Search time 124.4 seconds
(without alignments)
61.188 Million cell updates/sec

Title: US-09-641-104A-11

Perfect score: 227

Sequence: 1 NNYKNKMMVCQVGGIEALVR.....AGDREDITEPAICALRHLS 44

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	227	100.0	780	13	Q90424
2	227	100.0	781	11	Q9D335
3	227	100.0	781	13	Q42486
4	194	85.5	773	5	O76152
5	193	85.0	769	5	Q9NL44
6	181	79.7	806	5	Q25100
7	171	75.3	821	5	O61229
8	166	73.1	745	4	Q15151
9	166	73.1	745	4	Q9BWC4
10	159	70.0	745	11	P70565
11	148	65.2	729	13	Q9PVF7
12	131	57.7	678	5	O44326
13	74	32.6	811	5	Q18825
14	59.5	26.2	825	10	Q9LZA6
15	58.5	25.8	509	10	Q9STT1
16	58.5	25.8	533	10	Q9FJJ0

Q9Y446 homo sapien
Q9fm35 arabidopsis
Q9eq99 mus musculu
Q92dl6 listeria in
Q9c239 neurospora
Q9amk1 xanthomonas
Q9d988 morone saxa
Q91419 pseudomonas
Q87594 streptomyce
Q9tj58 streptomyce
Q99rw4 staphylococ
Q93l92 staphylococ
Q9erx3 rattus norv
Q93x72 gossypium h
Q92551 homo sapien
Q96n94 homo sapien
Q50202 rhodococcus
Q9uu1 schizosacch
Q9mls1 arabidopsis
Q91613 serratia ma
Q22948 caenorhabdi
Q96fs1 homo sapien
Q9u308 caenorhabdi
Q9ssj6 arabidopsis
Q9c9a6 arabidopsis
Q9kgj2 bacillus ha
Q24128 nicotiana t
Q81161 carica papa
Q9pez3 xylella fas

ALIGNMENTS

RESULT 1

Q90424
ID Q90424
AC Q90424
DT 01-NOV-1996 (TREMREL. 01, Created)
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)
DT 01-DEC-2001 (TREMREL. 19, Last annotation update)
DE B-CATENIN.
GN CTNNB.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Zebrafish;
OC Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96122902; PubMed=8562427;
RA Kelly G.M., Erezylmaz D.F., Moon R.T.;
RT "Induction of a secondary embryonic axis in zebrafish occurs following the overexpression of beta-catenin.";
RL Mech. Dev. 53:261-273(1995).
DR EMBL: U41081; AAC59732.1;
DR HSSP: P35222; IG3T
DR ZFIN: ZDB-GENE-980526-362; cttnb.
DR InterPro: IPR000225; Armadillo.
DR Pfam: PF00514; Armadillo_seg; 12.
DR SMART: SM00185; ARM; 11.
DR PROSITE: PS0176; ARM_REPEAT; 8.
SQ SEQUENCE 780 AA; 85542 MW; D7A1FB80F94066DC CRC64;

Query Match 100.0%; Score 227; DB 13; Length 780;
Best Local Similarity 100.0%; Pred. No. 1.5e-22;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNYKNKMMVCQVGGIEALVRTVLVRAGDREDITEPAICALRHLS 44
|||||
Db 429 NNYKNKMMVCQVGGIEALVRTVLVRAGDREDITEPAICALRHLS 472

```

RESULT 2
Q9D335 Q9D335 PRELIMINARY; PRT; 781 AA.
AC Q9D335
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE ADULT MALE COLON CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,
DE CLONE:9030417H18, FULL INSERT SEQUENCE.
GN CATNB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=COLON;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Schoenwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Willming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK018515; BAB31250.1; -.
DR HSSP; P35222; IG3J.
DR MGD; MGI:88276; Catnb.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 12.
DR SMART; SM00185; ARM; 12.
DR PROSITE; PS50176; ARM_REPEAT; 7.
DR SEQUENCE 781 AA; 85546 MW; 937538C3B5CD75D1 CRC64;

Query Match 100.0%; Score 227; DB 11; Length 781;
Best Local Similarity 100.0%; Pred. No. 1.5e-22;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNYKNKMMVCGVGIEALVTVLRAGDREDITEPAICALRHLS 44
Db 430 NNYKNKMMVCGVGIEALVTVLRAGDREDITEPAICALRHLS 473

RESULT 3
O42486 O42486 PRELIMINARY; PRT; 781 AA.
AC O42486;
DT 01-JAN-1998 (TEMBLrel. 05, Created)
DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE BETA CATENIN.
GN CHBCAT.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=WHITE LEGHORN; TISSUE=DORSAL SKIN;
RX MEDLINE=97464068; PubMed=9322759;
RA Ito J., Chuong C.M., Widelitz R.B.;
RT "Isolation and characterization of chicken beta-catenin.";
RL Gene 196:201-207(1997).
DR EMBL; U82964; AAB80856.1; -.
DR HSSP; P35222; IG3J.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 12.
DR SMART; SM00185; ARM; 11.
DR PROSITE; PS50176; ARM_REPEAT; 8.
DR SEQUENCE 781 AA; 85438 MW; 6D205D9A4DBAC562 CRC64;

Query Match 100.0%; Score 227; DB 13; Length 781;
Best Local Similarity 100.0%; Pred. No. 1.5e-22;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NNYKNKMMVCGVGIEALVTVLRAGDREDITEPAICALRHLS 44
Db 430 NNYKNKMMVCGVGIEALVTVLRAGDREDITEPAICALRHLS 473

RESULT 4
O76152 O76152 PRELIMINARY; PRT; 773 AA.
AC O76152;
DT 01-NOV-1998 (TEMBLrel. 08, Created)
DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE BETA-CATENIN.
OS Ciona savignyi.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;
OC Clonidae; Clonæ.
OX NCBI_TaxID=51511;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98443204; PubMed=9769178;
RA Yoshida S., Matikawa Y., Satoh N.;
RT "Regulation of the trunk-tail patterning in the ascidian embryo: a
RT possible interaction of cascades between lithium/beta-catenin and
RT localized maternal factor pem.";
RL Dev. Biol. 202:264-279(1998).
DR EMBL; AB012160; BAA32789.1; -.
DR HSSP; P35222; IG3J.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 11.
DR SMART; SM00185; ARM; 9.
DR PROSITE; PS50176; ARM_REPEAT; 7.
DR SEQUENCE 773 AA; 85217 MW; C1340CF82AFEBDAB CRC64;

Query Match 85.5%; Score 194; DB 5; Length 773;
Best Local Similarity 84.1%; Pred. No. 4.8e-18;
Matches 37; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 NNYKNKMMVCGVGIEALVTVLRAGDREDITEPAICALRHLS 44
Db 419 NNSNTRVCQVGIEALVTVLRAGDREDITEPSVICALRHLS 462

RESULT 5
O9NL44 O9NL44 PRELIMINARY; PRT; 769 AA.
AC O9NL44;
DT 01-OCT-2000 (TEMBLrel. 15, Created)
DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE BETA-CATENIN.
GN CIBETA-CATENIN.
OS Ciona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;

```

```

OC Cionidae; Ciona.
OX NCBI_TaxID=7719;
RN [1]
RP SEQUENCE FROM N.A.
RA Inai K., Takada N., Satoh N., Satou Y.;
RT "An essential role of beta-catenin in the endoderm specification of
RL ascidian embryo.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB031543; BAA92185.1; -
DR HSP: P35222; IG3J.
DR InterPro: IPR000225; Armadillo.
DR Pfam: PF00514; Armadillo_seg; 11.
DR SMART: SM00185; ARM; 11.
DR PROSITE: PS0176; ARM_REPEAT; 6.
SQ SEQUENCE 769 AA; 84703 MW; F61CC489B436E1BC CRC64;

Query Match 85.0%; Score 193; DB 5; Length 769;
Best Local Similarity 84.1%; Pred. No. 6.6e-18;
Matches 37; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 NNYKNKMMVCGVGIEALVRLVLRAGDREDITEPAICALRHLS 44
Db 420 NMSNKRVCQVGIEALVRLVLRAGDREDITEPTVCALRHLS 463

RESULT 6
Q25100 PRELIMINARY; PRT; 806 AA.
ID Q25100
AC Q25100;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE BETA-CATENIN
OS Hydra magnipapillata (Hydra).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;
OC Hydrozoa; Hydra.
OX NCBI_TaxID=6085;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WILDTYPE 105;
RX MEDLINE=96257271; PubMed=8654977;
RA Hohmayer E., Hatta M., Fischer R., Fujisawa T., Holstein T.W.,
RA Sugiyama T.;
RT "Identification of a Hydra homologue of the beta-
RT catenin/plakoglobin/armadillo gene family.";
RL Gene 172:155-159(1996).
DR EMBL: U36781; AAC47137.1; -
DR HSP: Q02248; 2BCT.
DR InterPro: IPR000225; Armadillo.
DR Pfam: PF00514; Armadillo_seg; 10.
DR SMART: SM00185; ARM; 10.
DR PROSITE: PS0176; ARM_REPEAT; 7.
SQ SEQUENCE 806 AA; 90462 MW; 689E5E982CD5051A CRC64;

Query Match 79.7%; Score 181; DB 5; Length 806;
Best Local Similarity 77.3%; Pred. No. 3.1e-16;
Matches 34; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 1 NNYKNKMMVCGVGIEALVRLVLRAGDREDITEPAICALRHLS 44
Db 489 NNPRKQVFGVGIEALVRLVLRAGDREDITEPAICALRHLS 532

RESULT 7
O61229 PRELIMINARY; PRT; 821 AA.
ID O61229
AC O61229;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE BETA-CATENIN.

```

```

OS Lytechinus variegatus (Sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidae; Euechinoidea; Echinacea; Temnopleuroidea; Toxopneustidae;
OC Lytechinus.
OX NCBI_TaxID=7654;
RN [1]
RP SEQUENCE FROM N.A.
RA Miller J.R., McClay D.R.;
RT "Changes in the pattern of adherens junction-associated beta-catenin
RT accompany morphogenesis in the sea urchin embryo.";
RL Dev. Biol. 192:310-322(1997).
DR EMBL: U34814; AAC06340.1; -
DR HSP: Q02248; 1DOW.
DR InterPro: IPR002086; Aldehyde_dehydr.
DR InterPro: IPR000225; Armadillo.
DR Pfam: PF00514; Armadillo_seg; 12.
DR SMART: SM00185; ARM; 12.
DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
DR PROSITE: PS0176; ARM_REPEAT; 8.
SQ SEQUENCE 821 AA; 89558 MW; 71E2ID562A99C5AD CRC64;

Query Match 75.3%; Score 171; DB 5; Length 821;
Best Local Similarity 68.2%; Pred. No. 7.3e-15;
Matches 30; Conservative 11; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NNYKNKMMVCGVGIEALVRLVLRAGDREDITEPAICALRHLS 44
Db 447 NNSRNMIVSQMAGVEALVQTLKAGDREITEPAICALRHVS 490

RESULT 8
Q15151 PRELIMINARY; PRT; 745 AA.
ID Q15151
AC Q15151; Q15093;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PLAKOGLOBIN.
GN JUP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89264555; PubMed=2726765;
RA Franke W.W., Goldschmidt M.D., Zimbelmann R., Mueller H.M.,
RA Schiller D.L., Cowin P.;
RT "Molecular cloning and amino acid sequence of human plakoglobin, the
RT common junctional plaque protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:4027-4031(1989).
RN [2]
RP SEQUENCE FROM N.A.
RA Zimbelmann R.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 239-409 FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=96157724; PubMed=8576101;
RA Ozawa M., Nuru K., Toyoyama H., Ohi Y.;
RT "Cloning of an alternative form of plakoglobin (gamma-catenin) lacking
RT the fourth armadillo repeat.";
RL J. Biochem. 118:836-840(1995).
RN [4]
RP SEQUENCE FROM N.A.
RX PubMed=11016852;
RA Whittock N.V., Eady R.A.J., McGrath J.A.;
RT "Genomic Organization and amplification of the human plakoglobin
RT gene.";
RL Exp. Dermatol. 9:323-326(2000).
DR EMBL: Z68228; CAA92522.1; -

```



```

RC STRAIN=N2;
RA Eisenmann D.M., Maloof J.N., Sims J.S., Keayon C., Kim S.K.;
RL Submitted (MAY1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: U046673; AAC48154.1; -
DR EMBL: AF063646; AAC17424.1; -
DR HSP: Q02248; 3bCT
DR TRANSFAC: T03880; -
SQ SEQUENCE 811 AA; 92227 MW; 97D6FFDE71BDDEFF CRC64;

Query: Match 32.6%; Score 74; DB 5; Length 811;
Best Local Similarity 38.9%; Pred. No. 0.14;
Matches 21; Conservative 3; Mismatches 18; Indels 12; Gaps 2;

Qy 1 NNYKNKMWCVGGIEALVFTVLR-----GD-----REDITEPAICALRHLL 42
Db 409 NNRHNKFCMNGWILVRALTATKEMGLDRKQAQMEDIYESLICTLRHL 462

RESULT 14
Q9LZA6 PRELIMINARY; PRT; 825 AA.
ID AC Q9LZA6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE HYPOTHEETICAL 92.1 KDA PROTEIN.
GN F8F6.250.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Peters S.A., van Staveren M., Dirkse W., Stiekema W.,
RA Bancroft I., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[2]
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RW EMBL: ALI62873; CAB85524.1; -
KW Hypothetical protein.
SQ SEQUENCE 825 AA; 92091 MW; A80717BD5971F144 CRC64;

Query Match 26.2%; Score 59.5; DB 10; Length 825;
Best Local Similarity 41.2%; Pred. No. 14;
Matches 14; Conservative 8; Mismatches 11; Indels 1; Gaps 1;

Qy 11 QVGGIEALVFTVLRAGDREDITEPAICALRHLLS 44
Db 298 QLGGVFSIVKRVQTGALHDIRQ-LQCLRLNLLS 330

RESULT 15
Q9STT1 PRELIMINARY; PRT; 509 AA.
ID AC Q9STT1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HYPOTHEETICAL 55.1 KDA PROTEIN.
GN T2J37.150.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Choise N., Robert C., Brottier P., Wincker P., Cattolico L.
```

RA Artiguenave F., Saurin W., Weissenbach J., Meves H.W., Mayer K.F.X.,
RA Lemcke K., Schueller C., Quetier F., Salanoubat M.,
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL049746; CAB41865.1; -;
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 5.
DR SMART; SM00185; ARM; 3.
DR SMART; SM00504; Ubox; 1.
DR PROSITE; PS50176; ARM_REPEAT; 1.
KW Hypothetical protein.
SQ SEQUENCE 509 AA; 55101 MW; 60F44BB1D1729B18 CRC64;

Query Match 25.8%; Score 58.5; DB 10; Length 509;
Best Local Similarity 35.0%; Pred. No. 12;
Matches 14; Conservative 10; Mismatches 15; Indels 1; Gaps 1;

Qy 4 KNKMMVCQGGIEALVRVLRAGDPREDITEPAICALRHLT 43
Db 298 ENKMVIGVIGAVEPLH-ALRSSESERARQDALALYHLS 336

Search completed: July 29, 2002, 16:12:11
Job time: 481 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 16:07:57 ; Search time 158.47 Seconds
(without alignments)
32.242 Million cell updates/sec

Title: us-09-641-104a-12

Perfect score: 241

Sequence: 1 RHQEAQAQNAVRLHYGLPV.....PPSHWPLIKATVGLIRNLAL 46

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

```

1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	241	100.0	781	21	Human beta catenin
2	241	100.0	781	21	Human beta-catenin
3	241	100.0	781	22	Novel human secret
4	241	100.0	781	22	Human beta-catenin
5	241	100.0	781	22	Mouse beta-catenin
6	241	100.0	800	22	Novel human secret
7	225.5	93.6	45	20	Human beta-catenin
8	195	80.9	840	22	Drosophila melanog
9	195	80.9	840	22	Drosophila melanog
10	195	80.9	840	22	Drosophila melanog
11	190	78.8	345	21	Human prostate can

12	190	78.8	700	12	AA11354	Cellular Receptor
13	158	24.1	749	19	AAW53457	Amino acid sequenc
14	157.5	23.9	205	22	ABB59041	Drosophila melanog
15	157.5	23.9	221	22	ABB57962	Drosophila melanog
16	155.5	23.0	167	22	AAW56559	Human brain expres
17	155.5	23.0	199	22	ABB59044	Drosophila melanog
18	154.5	22.6	127	22	AAO01902	Human polypeptide
19	154.5	22.6	192	19	AAW68475	HIV-1 strain YBF30
20	154.5	22.6	416	20	AAV02610	Amino acid sequenc
21	154.5	22.6	416	22	AAW98023	Pseudomonas mendoc
22	154.5	22.6	416	22	AAW84207	Amino acid sequenc
23	154	22.4	330	21	AAW53336	Human colon cancer
24	154	22.4	885	16	AAW76710	Recombinant cold-r
25	154	22.4	885	17	AAW90926	E. histolytica pyr
26	154	22.4	896	22	AAU53394	Propionibacterium
27	153.5	22.2	75	22	AAU60457	Propionibacterium
28	153	22.0	60	16	AAW81968	ORF encoded by exo
29	153	22.0	439	18	AAW17065	Aureobasidin resis
30	153	22.0	439	18	AAW17066	Aureobasidin sensi
31	153	22.0	439	19	AAW70519	Aspergillus nidula
32	153	22.0	439	22	AAW64146	Aspergillus nidula
33	153	22.0	634	22	ABG07516	Novel human diagno
34	153	22.0	890	22	ABG21377	Novel human diagno
35	153	22.0	932	21	AAW30697	Arabidopsis thalia
36	153	22.0	947	16	AAW81953	Rice pyruvate orth
37	153	22.0	953	16	AAW76707	Recombinant cold-r
38	153	22.0	953	17	AAW90923	F. bidentis pyruva
39	153	22.0	955	16	AAW76711	Recombinant cold-r
40	153	22.0	955	17	AAW90927	F. brownii cold re
41	153	22.0	957	21	AAW30696	Arabidopsis thalia
42	153	22.0	960	21	AAW30695	Arabidopsis thalia
43	152.5	21.8	1110	22	AAE03648	Human extracellular
44	152	21.6	79	22	ABW16565	Human nervous syst
45	152	21.6	227	22	ABG07377	Novel human diagno

ALIGNMENTS

RESULT 1
AAB07290
ID AAB07290 standard; Protein; 781 AA.
XX
AC AAE07290;
XX
DT 01-NOV-2000 (first entry)
XX
DE Human beta catenin.
XX
KW Beta catenin; cadherin; metastasis; cadherin-associated protein;
KW human; colorectal cancer; melanoma; antisense oligonucleotide;
KW gene therapy.
XX
OS Homo sapiens.
XX
PN US6066500-A.
XX
PD 23-MAY-2000.
XX
PF 25-JUN-1999; 99US-0344519.
XX
PR 25-JUN-1999; 99US-0344519.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Bennett CF, Cowser LM;
XX
DR WPI; 2000-410651/35.
XX
N-PSDB; AAA58320.
XX
PT New antisense compounds targeting nucleic acids encoding human beta
PT catenin (HBC) useful for treating diseases associated with HBC
PT expression and as prophylaxis to prevent or delay infection,

PT Inflammation or tumor formation -
 PS Example 13; Columns 45-52; 35pp; English.
 XX
 CC Beta catenin is a member of the catenin family of cytosolic proteins and
 CC a key member of the Wnt signalling pathway. Catenins interact with the
 CC cytoplasmic domains of cadherin glycoproteins, and are important in
 CC maintaining cell adhesiveness. The loss of cell adhesiveness is
 CC implicated in metastasis. Beta catenin is also known as
 CC cadherin-associated protein and is implicated in colorectal cancer and
 CC melanoma. The present sequence is the human beta catenin protein. The
 CC coding sequence of this protein was used in the present invention to
 CC design antisense oligonucleotides (AAA58327-A58366). The
 CC oligonucleotides are capable of hybridising to human beta catenin, in
 CC order to inhibit expression of human beta catenin. The oligonucleotides
 CC may be used in gene therapy for colorectal cancer or melanoma.
 XX
 SQ Sequence 781 AA;

Query Match 100.0%; Score 241; DB 21; Length 781;
 Best Local Similarity 100.0%; Pred. No. 1.6e-26;
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHQEAQAQNAVRLHYGLPVVVKLLHPPSHWPLIKATVGLIRNAL 46
 DB 474 rhqeaqaqnavrlhyglpvpvkvllhppshwplikatvglirnal 519

RESULT 2
 ID AAY70740 standard; protein; 781 AA.
 XX
 AC AAY70740;
 DT 24-JUL-2000 (first entry)
 XX
 DE Human beta-catenin.
 XX
 KW Human beta-catenin; Wnt antagonist; contraceptive; oocyte vaccine;
 KW oocyte development; female primate contraception; oocyte viability;
 KW monoclonal antibody; Wnt signalling.
 XX
 OS Homo sapiens.
 XX
 PN WO200021555-A1.
 XX
 PD 20-APR-2000.
 XX
 PF 13-OCT-1999; 99WO-US23640.
 XX
 PR 15-OCT-1998; 98US-0104355.
 XX
 PA (HARD) HARVARD COLLEGE.
 XX
 PI McMahon AP, Parr BA, Vaino S;
 XX
 DR WPI; 2000-317845/27.
 XX
 PT Contraceptive composition for inhibiting oocyte development in a female
 PT primate comprises a Wnt polypeptide antagonist -
 XX
 PS Example 3; Page 26; 57pp; English.
 XX

The patent discloses a method of female primate contraception comprising
 CC administering an antagonist of a Wnt polypeptide, inhibiting oocyte
 CC development. Wnt polypeptides are useful for promotive maturation of an
 CC immature oocyte. Wnt polypeptides are also useful for increasing the
 CC number of mature oocytes and to enhance oocyte viability. Soluble
 CC fragments of Wnt polypeptides have the ability to inhibit Wnt signalling,
 CC e.g., by blocking binding of a naturally-occurring Wnt protein to its
 CC receptor. They may be used to generate monoclonal antibodies which can
 CC inhibit oocyte development. The present sequence is the human

CC beta-catenin protein. Dominant negative mutants of beta-catenin lack one
 CC or more armadillo-like repeats which participate in cadherin binding.
 CC Other mutants include those lacking amino acids 555-781 or 424-781,
 CC or 1-422 of beta-catenin.
 XX
 SQ Sequence 781 AA;

Query Match 100.0%; Score 241; DB 21; Length 781;
 Best Local Similarity 100.0%; Pred. No. 1.6e-26;
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 01 RHQEAQAQNAVRLHYGLPVVVKLLHPPSHWPLIKATVGLIRNAL 46
 DB 474 rhqeaqaqnavrlhyglpvpvkvllhppshwplikatvglirnal 519

RESULT 3
 ID AAU28118 standard; Protein; 781 AA.
 XX
 AC AAU28118;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Novel human secretory protein, Seq ID No 287.
 XX
 KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
 KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
 KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
 KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
 KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
 KW gut protection; lung; liver fibrosis; immune deficiency; infection;
 KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
 KW fertility; analgesic; pain; antigen.
 XX
 OS Homo sapiens.
 XX
 PN WO200166689-A2;
 XX
 PD 13-SEP-2001.
 XX
 PF 05-MAR-2001; 2001WO-US04942.
 XX
 PR 07-MAR-2000; 2000US-0519705.
 PR 19-MAY-2000; 2000US-0574454.
 PR 17-JUN-2000; 2000US-0596193.
 PR 14-JUL-2000; 2000US-0616847.
 PR 19-SEP-2000; 2000US-0665363.
 PR 20-OCT-2000; 2000US-0693267.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
 PI Zheo QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
 XX
 DR WPI; 2001-589934/66.
 DR N-PSDB; AAS45018.
 XX
 PT Novel polypeptides and nucleic acids obtained from cDNA libraries
 PT prepared from various human tissues, for diagnosis and treatment of
 PT cancer, neurological, inflammatory, and autoimmune disorders -
 XX
 PS Example 4; SEQ ID No 287; 107pp; English.
 XX

The invention relates to novel isolated human secreted polypeptides (I)
 CC and polynucleotides (II). (I) and (II) are useful for treating
 CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
 CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
 CC involved in increasing haematopoiesis, stem cell survival, bone growth
 CC and remodeling. (I), (II) and modulators of (II) are useful for
 CC prophylaxis or treatment of one or more cancers. (II) is also useful for

creating transgenic animals useful for studying the in vivo activities of the polypeptide as well as for studying modulators of the polypeptides.
 (1) induces the proliferation of neural cells and regeneration of nerve and brain tissue and is useful for the treatment of central and peripheral nervous system diseases and neuropathies, such as Alzheimer's, Parkinson's disease, Huntington's disease, and amyotrophic lateral sclerosis. In addition, (1) is involved in chemotactic or chemokinetic activity, regulation of haematopoiesis and is useful for treating myeloid or lymphoid cell disorders, platelet disorders such as thrombocytopenia and for regeneration of bone, cartilage, tendon, ligament and/or nerve tissue growth, and in tissue repair, healing of burns, incisions, ulcers, for treating osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease. Furthermore, (1) is also useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, various immune deficiencies and disorders including severe combined immunodeficiency (SCID), bacterial or fungal infections, autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic reactions and conditions, such as asthma or other respiratory problems. In addition, (1) affects biorhythms or circadian cycles of rhythms, fertility, metabolism, catabolism, anabolism, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides analgesic effects or other pain reducing effects, immunoglobulin like activity and can act as an antigen in a vaccine composition to raise an immune response. AAU28020-AAU28395 represent novel human secreted protein amino acid sequences of the invention.

XX Sequence 781 AA;

*Query Match 100.0%; Score 241; DB 22; Length 781;
 Best Local Similarity 100.0%; Pred. No. 1.6e-26;
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHQAEAMQNAVRLHYGLPVVVKLLHPSPHWPLIKATVGLIRNLAL 46
 |||||
 Db 474 rhqaeamqnavrlhyglpvpvkvllhppshwplikatvgllrnlal 519

RESULT 4

AAE06038
 ID AAE06038 standard; Protein; 781 AA.

XX AC AAE06038;

XX DT 25-SEP-2001 (first entry)

XX DE Human beta-catenin protein #1.

XX KW Human; stem cell culturing; progenitor cell; pluripotential phenotype; transplantation; haematopoietic function; allogeneic recipient; signalling pathway; beta-catenin.

XX OS Homo sapiens.

XX PN WO200152649-A1.

XX PD 26-JUL-2001.

XX PF 17-JAN-2001; 2001WO-US01459.

XX PR 18-JAN-2000; 2000US-0176786.

XX PA (STRD) UNIV LELAND STANFORD JUNIOR.

XX PI Reya T, Nusse R, Weissman IL;

XX DR WPI; 2001-465328/50.

XX DR N-PSDB; AAD11164.

PT In vitro expansion of mammalian stem or progenitor cells, useful for producing cells that retain their pluripotential phenotype after expansion for use in transplantation, involves increasing the levels of

PT beta-catenin in the cell -

XX PS Disclosure; Page 23-24; 33pp; English.

XX CC The present invention relates to a method for in vitro expansion of mammalian stem or progenitor cells, which comprising increasing the intracellular concentration of beta-catenin in a progenitor or stem cell in an in vitro culture medium for a period sufficient for the progenitor or stem cell to divide. The number of cells having the functional phenotype of the stem or progenitor cells is expanded. The method is useful for culturing stem cells and progenitor cells, which retain their pluripotential phenotype after expansion, in vitro. The expanded cell populations are useful as a source of stem cells, e.g. to reconstitute function in a host that is deficient in a particular cell lineage or lineages. The expanded cell populations are also useful in transplantation to restore haematopoietic function to autologous or allogeneic recipients. The present sequence is human beta-catenin protein. Beta-catenin is a pivotal player in the signalling pathway initiated by Wnt proteins, which are mediators of several developmental processes.

XX SQ Sequence 781 AA;

Query Match 100.0%; Score 241; DB 22; Length 781;
 Best Local Similarity 100.0%; Pred. No. 1.6e-26;
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHQAEAMQNAVRLHYGLPVVVKLLHPSPHWPLIKATVGLIRNLAL 46
 |||||
 Db 474 rhqaeamqnavrlhyglpvpvkvllhppshwplikatvgllrnlal 519

RESULT 5

AAE06039
 ID AAE06039 standard; Protein; 781 AA.

XX AC AAE06039;

XX DT 25-SEP-2001 (first entry)

XX DE Mouse beta-catenin protein #2.

XX KW Mouse; stem cell culturing; progenitor cell; pluripotential phenotype; transplantation; haematopoietic function; allogeneic recipient; signalling pathway; beta-catenin.

XX OS Mus musculus.

XX PN WO200152649-A1.

XX PD 26-JUL-2001.

XX PF 17-JAN-2001; 2001WO-US01459.

XX PR 18-JAN-2000; 2000US-0176786.

XX PA (STRD) UNIV LELAND STANFORD JUNIOR.

XX PI Reya T, Nusse R, Weissman IL;

XX DR WPI; 2001-465328/50.

XX DR N-PSDB; AAD11165.

PT In vitro expansion of mammalian stem or progenitor cells, useful for producing cells that retain their pluripotential phenotype after expansion for use in transplantation, involves increasing the levels of beta-catenin in the cell -

XX PS Disclosure; Page 28-30; 33pp; English.

XX CC The present invention relates to a method for in vitro expansion of mammalian stem or progenitor cells, which comprising increasing the

CC intracellular concentration of beta-catenin in a progenitor or stem cell
 CC in an in vitro culture medium for a period sufficient for the progenitor
 CC or stem cell to divide. The number of cells having the functional
 CC phenotype of the stem or progenitor cells is expanded. The method is
 CC useful for culturing stem cells and progenitor cells, which retain their
 CC pluripotential phenotype after expansion, in vitro. The expanded cell
 CC populations are useful as a source of stem cells, e.g. to reconstitute
 CC function in a host that is deficient in a particular cell lineage or
 CC lineages. The expanded cell populations are also useful in
 CC transplantation to restore haematopoietic function to autologous or
 CC allogeneic recipients. The present sequence is mouse beta-catenin
 CC protein. Beta-catenin is a pivotal player in the signalling pathway
 CC initiated by Wnt proteins, which are mediators of several developmental
 CC processes.

XX Sequence 781 AA;

Query Match 100.0%; Score 241; DB 22; Length 781;
 Best Local Similarity 100.0%; Pred. No. 1.6e-26;
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHQAEAMQNAVRHYGLPVVVKLLHPPSHWPLIKATVGLIRNLAL 46

Db 474 rhqaeamqnavrhyglpvpvkvllhppshwplikatvglirnlal 519

RESULT 6

AAU28306
 ID AAU28306 standard; Protein: 800 AA.

XX AC AAU28306;

DT 18-DEC-2001 (first entry)

XX Novel human secretory protein, Seq ID No 663.

XX Human: secreted protein; arthritis; Crohn's disease; sepsis; shock;
 KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
 KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
 KW anyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
 KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
 KW gut protection; lung; liver fibrosis; immune deficiency; infection;
 KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
 KW fertility; analgesic; pain; antigen.

XX Homo sapiens.

XX WO200166689-A2.

XX 13-SEP-2001.

XX 05-MAR-2001; 2001WO-US04942.

XX 07-MAR-2000; 2000US-0519705.

XX 19-MAY-2000; 2000US-0574454.

XX 17-JUN-2000; 2000US-0596193.

XX 14-JUL-2000; 2000US-0616847.

XX 19-SEP-2000; 2000US-0665363.

XX 20-OCT-2000; 2000US-0693267.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;

PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;

XX WPI; 2001-589934/66.

XX N-PSDB; AAS45206.

XX Novel polypeptides and nucleic acids obtained from cDNA libraries
 PT prepared from various human tissues, for diagnosis and treatment of
 PT cancer, neurological, inflammatory, and autoimmune disorders

XX Example 2; SEQ.ID.No.663; 107pp; English.

XX This invention relates to novel isolated human secreted polypeptides (I)
 CC and polynucleotides (II). (I) and (II) are useful for treating
 CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
 CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
 CC involved in increasing haematopoiesis, stem cell survival, bone growth
 CC and remodeling. (I), (II) and modulators of (I) are useful for
 CC prophylaxis or treatment of one or more cancers. (II) is also useful for
 CC creating transgenic animals useful for studying the in vivo activities of
 CC the polypeptide as well as for studying modulators of the polypeptides.
 CC (I) induces the proliferation of neural cells and regeneration of nerve
 CC and brain tissue and is useful for the treatment of central and
 CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
 CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
 CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
 CC activity, regulation of haematopoiesis and is useful for treating myeloid
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 CC tissue growth, and in tissue repair, healing of burns, incisions,
 CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
 CC disorders, or periodontal disease. Furthermore, (I) is also useful for
 CC gut protection, or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues, various immune deficiencies and
 CC disorders including severe combined immunodeficiency (SCID), bacterial or
 CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
 CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
 CC reactions and conditions, such as asthma or other respiratory problems.
 CC In addition, (I) affects biorhythms or circadian cycles of rhythms, of
 CC fertility, metabolism, catabolism, anabolism, storage or elimination of
 CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
 CC analgesic effects or other pain reducing effects, immunoglobulin like
 CC activity and can act as an antigen in a vaccine composition to raise an
 CC immune response. AAU28020-AAU28395 represent novel human secreted protein
 CC amino acid sequences of the invention.

XX SQ Sequence 800 AA;

Query Match 100.0%; Score 241; DB 22; Length 800;

Best Local Similarity 100.0%; Pred. No. 1.7e-26;

Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHQAEAMQNAVRHYGLPVVVKLLHPPSHWPLIKATVGLIRNLAL 46

Db 492 rhqaeamqnavrhyglpvpvkvllhppshwplikatvglirnlal 537

RESULT 7

AAV33238

ID AAV33238 standard; peptide: 45 AA.

XX AC AAV33238;

XX 18-NOV-1999 (first entry)

XX Human: beta-catenin-protein mutant armadillo repeat arm 9.

XX Beta-catenin; human; armadillo repeat; treatment; human disease; LEF-1;
 KW modulator; interaction domain; transcription factor; TCF-4; oncogenic;
 KW tumor suppressor gene; conductin; E-cadherin; screening; carcinoma; APC;
 KW Wnt signaling pathway; tumor development; anti-oncogenic; melanoma;
 KW organ regeneration; tissue regeneration; hair growth; mutant.

XX Homo sapiens.

XX Synthetic.

XX DE19909251-A1.

XX 26-AUG-1999.

XX 22-FEB-1999; 99DE-1009251.

XX PR 21-FEB-1998; 98DE-1007390.
XX PA (DELB-) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.
XX PI Birchmeier W, Von Kries J;
XX DR WPI; 1999-470389/40.
XX Agents for treating human diseases, particularly cancer, modulate
PT interaction of beta-catenin with transcription factors or tumor
PT suppressor gene products
XX Example 4; Fig 5; 16pp; German.
XX This invention describes a novel agent (A) for treating human disease
CC which is based on substances (I) that modulate (inhibit or promote) the
CC interaction of beta-catenin with transcription factors or products of
CC tumor suppressor genes. The invention also describes (a) peptides (II)
CC comprising part of the LEF-1/TCF-4 transcription factors, or their
CC variants and mutants; (b) peptides and related molecules (III) from the
CC armadillo domain (arm units 3-8) of beta-catenin, and mutants of the
CC complete beta-catenin molecule, that include at least one of the
CC specific interaction domains for LEF-1, TCF-4, APC, conductin or
CC E-cadherin; and (c) enzyme-linked immunosorbent assay (ELISA) for
CC screening substance libraries for compounds that modulate interaction of
CC beta-catenin with LEF-1/TCF, APC, conductin or E-cadherin. Beta-catenin
CC is a key compound in the Wnt signalling pathway and is involved in
CC development of tumors. Generally its interaction with LEF-1 or TCF-4 is
CC oncogenic but interaction with APC, conductin or E-cadherin is
CC anti-oncogenic. (A) which inhibit interaction are particularly used to
CC treat tumors, especially carcinoma of the colon and melanoma, but also,
CC where they promote interaction, to stimulate regeneration of organs and
CC tissues, specifically hair growth. AAV3230-Y33241 represent mutant
CC human beta-catenin armadillo repeat fragments described in the method of
CC the invention.
XX
XX SQ Sequence 45 AA;

Query Match 93.6%; Score 225.5; DB 20; Length 45;
Best Local Similarity 97.8%; Pred. NO. 9.6e-26;
Matches 45; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 RHQEAEMQNAVRLHYGLPVVVKLLHPPSHWPLIRATVGLIRNAL 46
|||||
Db 1 rhqeaemaqnavrlhyglpvvvklhppshwpli-atvglirnal 45

RESULT 8

ABB60196
ID ABB60196 standard; Protein; 840 AA.
XX
AC ABB60196;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 7380.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR Drosophila; developmental biology; cell signalling; insecticide;
pharmaceutical.
KW
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX

PA (PEXE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL04299.
XX

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions
XX
PS Disclosure; SEQ ID NO 7380; 21pp + Sequence Listing; English.
XX

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 840 AA;

Query Match 80.9%; Score 195; DB 22; Length 840;
Best Local Similarity 78.3%; Pred. NO. 1e-19;
Matches 36; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 RHQEAEMQNAVRLHYGLPVVVKLLHPPSHWPLIKATVGLIRNAL 46
|||||
Db 479 rhvdselaqnavrlhyglsvivklhppsrwplikavglirnal 524

RESULT 9

ABB65819.
ID ABB65819 standard; Protein; 840 AA.
XX
AC ABB65819;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 24249.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEXE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL09922.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions
XX
PS Disclosure; SEQ ID NO 24249; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 840 AA;

Query Match 80.9%; Score 195; DB 22; Length 840;
Best Local Similarity 78.3%; Pred. No. 1e-19;
Matches 36; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
QY 1 RHQEAEMQNAVRLHYGLPVVVKLLHPPSHWPLIKATVGLIRNAL 46
DB 479 rhvdselaqnavrlngysvkvllhppsrlpwlikavglirnal 524

RESULT 10
ABB65821
ID ABB65821 standard; Protein; 840 AA.
XX AC ABB65821;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster polypeptide SEQ ID NO 24255.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-0509231.
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
PI WIPI; 2001-656860/75.
PI N-PSDB; ABL05924.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions.
XX PS Disclosure; SEQ ID NO 24255; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABB57737-ABB72072).
XX CC The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 840 AA;
Query Match 80.9%; Score 195; DB 22; Length 840;
Best Local Similarity 78.3%; Pred. No. 1e-19;
Matches 36; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
QY 1 RHQEAEMQNAVRLHYGLPVVVKLLHPPSHWPLIKATVGLIRNAL 46
DB 479 rhvdselaqnavrlngysvkvllhppsrlpwlikavglirnal 524

RESULT 11
AAB56810
ID AAB56810 standard; Protein; 345 AA.
XX AC AAB56810;
XX DT 13-MAR-2001 (first entry)
XX DE Human prostate cancer antigen protein sequence SEQ ID NO:1388.
XX KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
XX KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
XX KW vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
XX KW antibacterial; gene therapy; neural; immune; reproductive; renal;
XX KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
XX KW wound; infectious disease.
XX OS Homo sapiens.
XX PN WO200055174-A1.
XX PD 21-SEP-2000.
XX PF 08-MAR-2000; 2000WO-0505988.
XX PR 12-MAR-1999; 99US-0124270.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PA (ROSE/) ROSEN C A.
XX PI Rosen, CA, Ruben SM;
XX WIPI; 2000-587513/55.
XX N-PSDB; AAF16013.
XX PT Prostate cancer associated gene sequences, referred to as prostate
XX PT cancer antigens, useful for treatment, prevention, and diagnosis of
XX PT disorders such as prostate cancer.
XX PS Claim 11; Page 1813-1815; 2338pp; English.

XX AAF15566 to AAF16505 encode the human prostate cancer associated
XX proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
XX CC The prostate cancer antigens can have neuroprotective, cytostatic,
XX CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
XX CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
XX CC and can be used in gene therapy. The prostate cancer antigen
XX CC polynucleotides may be used for detection of prostate cancer, chromosome
XX CC identification, as chromosome markers, and for numerous other diagnostic
XX CC or research purposes. The prostate cancer antigens may be used to treat
XX CC disorders such as neural, immune, muscular, reproductive,
XX CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
XX CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
XX CC AAB57303 represent sequences used in the exemplification of the present
XX CC invention.

XX SQ Sequence 345 AA;

Query Match 78.8%; Score 190; DB 21; Length 345;

```

Matches: 34; Conservative 8; Mismatches 4; Indels 0; Gaps 0;
QY 11 RHOAEAMQAVRLHYGLPVVVKLLHPPSHWPLIKATVGLIRNAL 46
      |||||:||||:||||:||||:||||:||||:||||:||||:
Db 422 rhoaeamagnsrvlvdgkavikvllngdngdplvkatiglrinal 467

```

XX	AAW53457	
XX	AAW53457 standard; Protein; 749 AA.	
XX	AC	
XX	AAW53457;	
XX	28-JUL-1998 (first entry)	
XX	Amino acid sequence of human tumour PEX protein.	
DE		
XX	Recombinant tumour PEX-protein; phosphate regulating gene: inhibition;	
KW	endopeptidases; metalloendopeptidase; tumour; OHO; PEX active site;	
KW	oncogenous hypophosphataemia osteomalacia; chronic renal failure;	
KW	idiopathic hypercalcuria.	
XX		
OS	Homo sapiens.	
XX		
PX	WOS810078-A2.	
PN		
XX		
PD	12-MAR-1998.	
XX		
PF	04-SEP-1997; 97WO-CA00617.	
XX		
PR	05-SEP-1996; 96US-0025454.	
XX		
PA	(GOLT/) GOLTZMAN D.	
PA	(HEND/) HENDERSON J E.	
PA	(KARA/) KARAPLIS A C.	
PA	(LIPM/) LIPMAN M L.	
PA	(PAND/) PANDA D.	
PA	(SHEN/) SHEN Y.	
XX		
PI	Goltzman D, Henderson JE, Karaplis AC, Lipman ML;	
PI	Panda D, Shen Y;	
XX		
DR	WPI; 1998-193628/17.	
DR	N-PSDB; AAV23264.	
XX		
PT	DNA encoding human PEX - useful for the treatment of	
PT	hyperphosphataemia or chronic renal failure	
XX		
PS	Disclosure; Fig 1A-1G; 27pp; English.	
XX		
CC	This is the amino acid sequence of the recombinant human tumour PEX	
CC	protein a product of a phosphate regulating gene with homology to	
CC	endopeptidases, suggesting that PEX might be a metalloendopeptidase.	
CC	It was isolated from tumours causing oncogenous hypophosphataemia	
CC	osteomalacia (OHO). Compounds which are targeted to inhibit PEX	
CC	active site, identified in the method of the invention, are useful	
CC	for the treatment of hyperphosphataemia or chronic renal failure.	
CC	PEX can also be used as a target for the treatment of idiopathic	
XX	hypercalcuria.	
XX		
SQ	Sequence 749 AA:	

```

Query Match      24.1%; Score 58; DB 19; Length 749;
Best Local Similarity 30.3%; Pred. No. 11;
Matches 10; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

OY      7 MAQNAVRLHYGLPVVVKLLHPSSHPLIKATVG 39
      | | | | | | | | | | | | | | | |
Db      143 mnekaiekadgkplhlhrhsfwrpvniesnig 175

RESULT 14

```

ABB59041
ID ABB59041 standard; Protein; 205 AA.

XX AC ABB59041;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 3915.

XX KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US09231.

XX PR 23-MAR-2000; 2000US-191637P.

XX PR 11-JUL-2000; 2000US-0614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX DR WPI; 2001-656860/75.

XX DR N-PSDB; ABL03144.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -

XX PS Disclosure; SEQ ID NO 3915; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072).

XX CC The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 205 AA;

Query Match 23.9%; Score 57.5; DB 22; Length 205;

Best Local Similarity 37.2%; Pred. No. 2.7;

Matches 16; Conservative 7; Mismatches 17; Indels 3; Gaps 2;

QY 5 AEMAQNAVRLHYGLPVVVKLLHPPSHW--PLIKATVGLIRNLA 45

Db 141 aqyaapav-ahyaapavvktvavahyaapavvktvavahya 182

RESULT 15

ABB57962

ID ABB57962 standard; Protein; 221 AA.

XX AC ABB57962;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 678.

XX KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.

XX

OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US09231.

XX PR 23-MAR-2000; 2000US-191637P.

XX PR 11-JUL-2000; 2000US-0614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX DR WPI; 2001-656860/75.

XX DR N-PSDB; ABL02065.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -

XX PS Disclosure; SEQ ID NO 678; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072).

XX CC The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 221 AA;

Query Match 23.9%; Score 57.5; DB 22; Length 221;

Best Local Similarity 37.2%; Pred. No. 2.9;

Matches 16; Conservative 7; Mismatches 17; Indels 3; Gaps 2;

QY 5 AEMAQNAVRLHYGLPVVVKLLHPPSHW--PLIKATVGLIRNLA 45

Db 141 aqyaapav-ahyaapavvktvavahyaapavvktvavahya 182

Search completed: July 29, 2002, 16:07:58

Job time: 648 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 16:05:12 ; Search time 57.6 Seconds
(without alignments)

19,507 Million cell updates/sec

Title: US-09-641-104A-12

Perfect score: 241

Sequence: 1 RHQEAEMAGNAVRLHYGLPV.....PPSHWPLIKATVGLIRNAL 46

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	54.5	22.6	416	4	US-09-333-208-2
2	54.5	22.6	416	4	US-09-333-254-2
3	54	22.4	885	2	US-08-500-857A-8
4	53	22.0	439	4	US-08-716-873-2
5	53	22.0	439	4	US-08-716-873-4
6	53	22.0	439	4	US-09-368-431-2
7	53	22.0	439	4	US-09-368-431-4
8	53	22.0	953	2	US-08-500-857A-2
9	53	22.0	955	2	US-08-500-857A-10
10	51	21.2	649	4	US-09-300-909-15
11	51	21.2	649	4	US-09-300-909-27
12	51	21.2	878	3	US-08-941-936-2
13	49.5	20.5	405	4	US-09-608-285A-25
14	49.5	20.5	428	4	US-09-608-285A-3
15	49.5	20.5	428	4	US-09-608-285A-5
16	49.5	20.5	428	4	US-09-608-285A-7
17	49.5	20.5	428	4	US-09-240-639-6
18	49.5	20.5	428	4	US-09-240-639-9
19	49	20.3	374	2	US-09-007-438-2
20	49	20.3	374	3	US-09-207-493-2
21	49	20.3	419	3	US-09-335-409-8
22	49	20.3	419	4	US-09-413-814-71
23	49	20.3	419	4	US-09-568-102-8
24	49	20.3	419	4	US-09-567-969-8
25	49	20.3	419	4	US-09-568-480-8
26	49	20.3	419	4	US-09-568-486-8
27	49	20.3	419	4	US-09-568-472-8

28 49 20.3 436 4 US-08-716-873-5 Sequence 5, Appli
29 49 20.3 436 4 US-09-368-431-5 Sequence 5, Appli
30 48.5 20.1 406 1 US-08-155-906-2 Sequence 2, Appli
31 48 19.9 1829 4 US-09-157-420-1 Sequence 1, Appli
32 47 19.5 514 4 US-08-796-899-25 Sequence 25, Appli
33 46.5 19.3 341 3 US-08-890-719-38 Sequence 38, Appli
34 46.5 19.3 365 3 US-08-652-265-23 Sequence 23, Appli
35 46.5 19.3 365 4 US-08-834-497A-23 Sequence 23, Appli
36 46.5 19.3 365 4 US-09-503-444A-23 Sequence 9, Appli
37 46.5 19.3 1150 4 US-09-238-303-9 Sequence 13, Appli
38 46.5 19.3 3665 2 US-08-222-617A-13 Sequence 4, Appli
39 46.5 19.3 3712 2 US-08-222-617A-25 Sequence 25, Appli
40 46.5 19.3 3712 2 US-08-222-617A-25 Sequence 25, Appli
41 46 19.1 219 1 US-08-470-261-2 Sequence 2, Appli
42 46 19.1 219 3 US-08-916-989B-2 Sequence 2, Appli
43 46 19.1 219 4 US-09-432-253-2 Sequence 2, Appli
44 46 19.1 219 5 PCT-US94-13187-2 Sequence 2, Appli
45 46 19.1 1612 3 US-08-545-860D-48 Sequence 48, Appli

ALIGNMENTS

RESULT 1

US-09-333-208-2

; Sequence 2, Application US/09333208A

; Patent No. 6225089

; GENERAL INFORMATION:

; APPLICANT: Chen, Kevin K.

; TITLE OF INVENTION: A Putative Solvent/Antibiotic Resistant Gene from

; FILE OF INVENTION: Pseudomonas Mendocina

; FILE REFERENCE: CL-1160-C

; CURRENT APPLICATION NUMBER: US/09/333,208A

; CURRENT FILING DATE: 1999-06-15

; EARLIER APPLICATION NUMBER: 09/183,270

; EARLIER FILING DATE: 1998-10-30

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 2

; LENGTH: 416

; TYPE: PRT

; ORGANISM: Pseudomonas mendocina KR-1

US-09-333-208-2

Query Match 22.6%; Score 54.5; DB 4; Length 416;
Best Local Similarity 38.7%; Pred. No. 3.9; Mismatches 12; Indels 3; Gaps 1;
Matches 12; Conservative 4;

QY 11 AVRLHYGLPVVVKLL--HPPSHWPLIKATV 38

DB 344 AMLYYGVVTAALIFVQPSGHWPLVAAVL 374

RESULT 2

US-09-333-254-2

; Sequence 2, Application US/09333254A

; Patent No. 6235882

; GENERAL INFORMATION:

; APPLICANT: Chen, Kevin K.

; TITLE OF INVENTION: A Gene Encoding a Putative Efflux Protein for Solvents/

; FILE OF INVENTION: Antibiotics in Pseudomonas Mendocina

; FILE REFERENCE: CL-1160-D

; CURRENT APPLICATION NUMBER: US/09/333,254A

; CURRENT FILING DATE: 1999-06-15

; EARLIER APPLICATION NUMBER: 08/961,738

; EARLIER FILING DATE: 1997-10-31

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 2

; LENGTH: 416

; TYPE: PRT

; ORGANISM: Pseudomonas mendocina KR-1

;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Warren M. Cheek, Jr.
;; REGISTRATION NUMBER: 33,367
;; REFERENCE/DOCKET NUMBER:
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-371-8850
;; TELEFAX:
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 439 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; US-08-716-873-4

Query Match 22.0%; Score 53; DB 4; Length 439;
Best Local Similarity 38.9%; Pred. No. 7;
Matches 14; Conservative 2; Mismatches 16; Indels 1;

Qy 15 HYGLPVVVKLL----HPPSHWPLIKATVGLIRNLAL 46
||| | | | : || | | | | :
Db 192 HYGAPFVCSLIMFIFGPGTVPFLFARTFGYISMTAV 227

RESULT 6
US-09-368-431-2
;; Sequence 2, Application US/09368431
;; Patent No. 6294651
;; GENERAL INFORMATION:
;; APPLICANT: Takashi OKADO et al.
;; TITLE OF INVENTION: GENE REGULATING AUREOBASIDIN SENSITIVITY
;; TITLE OF INVENTION: (AS AMENDED)
;; NUMBER OF SEQUENCES: 50
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: WENDEROTH, LIND & PONACK, L.L.P.
;; STREET: 2033 K Street, N.W., #800
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: U.S.A.
;; ZIP: 20006
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: MS-DOS
;; SOFTWARE: Wordperfect 5.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/368,431
;; FILING DATE: August 5, 1999
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/716,873
;; FILING DATE: September 20, 1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Warren M. Cheek, Jr.
;; REGISTRATION NUMBER: 33,367
;; REFERENCE/DOCKET NUMBER:
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-721-8200
;; TELEFAX: 202-721-8250
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 439 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; US-09-368-431-2

Query Match 22.0%; Score 53; DB 4; Length 439;
Best Local Similarity 38.9%; Pred. No. 7;
Matches 14; Conservative 2; Mismatches 16; Indels 1;

Qy 15 HYGLPVVVKLL----HPPSHWPLIKATVGLIRNLAL 46
||| | | | : || | | | | :
Db 192 HYGAPFVCSLIMFIFGPGTVPFLFARTFGYISMTAV 227

RESULT 7
US-09-368-431-4
;; Sequence 4, Application US/09368431
;; Patent No. 6294651
;; GENERAL INFORMATION:
;; APPLICANT: Takashi OKADO et al.
;; TITLE OF INVENTION: GENE REGULATING AUREOBASIDIN SENSITIVITY
;; TITLE OF INVENTION: (AS AMENDED)
;; NUMBER OF SEQUENCES: 50
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: WENDEROTH, LIND & PONACK, L.L.P.
;; STREET: 2033 K Street, N.W., #800
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: U.S.A.
;; ZIP: 20006
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: MS-DOS
;; SOFTWARE: Wordperfect 5.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/368,431
;; FILING DATE: August 5, 1999
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/716,873
;; FILING DATE: September 20, 1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Warren M. Cheek, Jr.
;; REGISTRATION NUMBER: 33,367
;; REFERENCE/DOCKET NUMBER:
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-721-8200
;; TELEFAX: 202-721-8250
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 439 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; US-09-368-431-4

Query Match 22.0%; Score 53; DB 4; Length 439;
Best Local Similarity 38.9%; Pred. No. 7;
Matches 14; Conservative 2; Mismatches 16; Indels 1;

Qy 15 HYGLPVVVKLL----HPPSHWPLIKATVGLIRNLAL 46
||| | | | : || | | | | :
Db 192 HYGAPFVCSLIMFIFGPGTVPFLFARTFGYISMTAV 227

RESULT 8
US-08-500-857A-2
;; Sequence 2, Application US/08500857A
;; Patent No. 5912156
;; GENERAL INFORMATION:
;; APPLICANT: OHATA, SHOZO
;; APPLICANT: USAMI, STORU
;; APPLICANT: BURNELL, JAMES N

```

; TITLE OF INVENTION: POLYPEPTIDE HAVING COLD-STABLE PYRUVATE,
; TITLE OF INVENTION: ORTHOPHOSPHATE KINASE ACTIVITY, DNA ENCODING THE SAME
; TITLE OF INVENTION: AND RECOMBINANT VECTOR AND TRANSFORMED PLANTS CONTAINING
; TITLE OF INVENTION: THE DNA
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
; STREET: 8110 GATE HOUSE ROAD SUITE 500 EAST
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/500,857A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURPHY JR, GERALD M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 760-208P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 103-205-8050
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 953 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-500-857A-2

Query Match 22.0%; Score 53; DB 2; Length 953;
Best Local Similarity 64.3%; Pred. No. 18;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 17 GLPVVVKLLHPPSH 30
DB 692 GLPVTIRLLDPLH 705

RESULT 9
US-08-500-857A-10
; Sequence 10, Application US/08500857A
; Patent No. 5912156
; GENERAL INFORMATION:
; APPLICANT: OHATA, SHOZO
; APPLICANT: USAMI, STORU
; APPLICANT: BURNELL, JAMES N
; TITLE OF INVENTION: POLYPEPTIDE HAVING COLD-STABLE PYRUVATE,
; TITLE OF INVENTION: ORTHOPHOSPHATE KINASE ACTIVITY, DNA ENCODING THE SAME
; TITLE OF INVENTION: AND RECOMBINANT VECTOR AND TRANSFORMED PLANTS CONTAINING
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
; STREET: 8110 GATE HOUSE ROAD SUITE 500 EAST
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/500,857A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURPHY JR, GERALD M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 760-208P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 103-205-8050
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 953 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-500-857A-2

Query Match 22.0%; Score 53; DB 2; Length 953;
Best Local Similarity 64.3%; Pred. No. 18;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 17 GLPVVVKLLHPPSH 30
DB 692 GLPVTIRLLDPLH 705

RESULT 9
US-08-500-857A-10
; Sequence 10, Application US/08500857A
; Patent No. 5912156
; GENERAL INFORMATION:
; APPLICANT: OHATA, SHOZO
; APPLICANT: USAMI, STORU
; APPLICANT: BURNELL, JAMES N
; TITLE OF INVENTION: POLYPEPTIDE HAVING COLD-STABLE PYRUVATE,
; TITLE OF INVENTION: ORTHOPHOSPHATE KINASE ACTIVITY, DNA ENCODING THE SAME
; TITLE OF INVENTION: AND RECOMBINANT VECTOR AND TRANSFORMED PLANTS CONTAINING
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
; STREET: 8110 GATE HOUSE ROAD SUITE 500 EAST
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/500,857A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURPHY JR, GERALD M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 760-208P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 103-205-8050
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 953 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-500-857A-2

Query Match 22.0%; Score 53; DB 2; Length 955;
Best Local Similarity 64.3%; Pred. No. 18;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 17 GLPVVVKLLHPPSH 30
DB 694 GLPVTIRLLDPLH 707

RESULT 10
US-09-300-909-15
; Sequence 15, Application US/09300909
; Patent No. 6306580
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PREPARATION OF HUMAN PAPILLOMAVIRUS E1 HAVING
; TITLE OF INVENTION: HELICASE ACTIVITY AND METHOD THEREFOR
; NUMBER OF SEQUENCES: 27
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/300,909
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/083,942
; FILING DATE: 01-MAY-1998
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 649 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-300-909-15

Query Match 21.2%; Score 51; DB 4; Length 649;
Best Local Similarity 45.8%; Pred. No. 22;
Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 176 HQEAEAPNSVREHAGTAGILELL 199
DB 176 HQEAEAPNSVREHAGTAGILELL 199

RESULT 11
US-09-300-909-27
; Sequence 27, Application US/09300909
; Patent No. 6306580
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PREPARATION OF HUMAN PAPILLOMAVIRUS E1 HAVING
```

;; TITLE OF INVENTION: HELICASE ACTIVITY AND METHOD THEREFOR
;; NUMBER OF SEQUENCES: 27
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/300,909
;; FILING DATE:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 60/083,942
;; FILING DATE: 01-MAY-1998
;; INFORMATION FOR SEQ ID NO: 27:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 649 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-09-300-909-27

Query Match 21.2%; Score 51; DB 4; Length 649;
Best Local Similarity 45.8%; Pred. No. 22;
Matches 11; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 2 HQEAMQNAVRLHYGLPVVVKLL 25
| | | | | : | | | | : | | | |
Db 176 HTEAAPTNSVREHAGTIGLELL 199

RESULT 12
US-08-941-936-2
;; Sequence 2, Application US/08941936
;; Patent No. 6054305
;; GENERAL INFORMATION:
;; APPLICANT: Tatsumi, Hiroki
;; APPLICANT: Eisaki, Naoki
;; APPLICANT: Horiuchi, Tatsuo
;; APPLICANT: Nagahara, Ayumu
;; TITLE OF INVENTION: Pyruvate Orthophosphate Dikinase Gene,
;; TITLE OF INVENTION: Recombinant DNA, and Process For Producing Pyruvate
;; TITLE OF INVENTION: Orthophosphate Dikinase
;; NUMBER OF SEQUENCES: 8
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: MEDLEN & CARROLL, LLP
;; STREET: 220 Montgomery Street, Suite 2200
;; CITY: San Francisco
;; STATE: CA
;; COUNTRY: US
;; ZIP: 94104
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/941,936
;; FILING DATE: 01-OCT-1997
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Carroll, Peter G.
;; REGISTRATION NUMBER: 32,837
;; REFERENCE/DOCKET NUMBER: HIRAKI-03009
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415-705-8410
;; TELEFAX: 415-397-8338
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 878 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear

;; MOLECULE TYPE: protein
;; US-08-941-936-2
Query Match 21.2%; Score 51; DB 3; Length 878;
Best Local Similarity 57.1%; Pred. No. 32;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 17 GLPVVVVKLLHPPSH 30
| | | | | : | | | | : | | | |
Db 617 GLPVTIRLIDPPLH 630

RESULT 13
US-09-608-285A-25
;; Sequence 25, Application US/09608285A
;; Patent No. 6335013
;; GENERAL INFORMATION:
;; APPLICANT: Ford, John
;; APPLICANT: Mulero, Julio
;; APPLICANT: Yeung, George
;; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
;; TITLE OF INVENTION: POLYPEPTIDES
;; FILE REFERENCE: 28110/36570
;; CURRENT APPLICATION NUMBER: US/09/608,285A
;; CURRENT FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: 09/583,231
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 09/557,800
;; PRIOR FILING DATE: 2000-04-25
;; PRIOR APPLICATION NUMBER: 09/481,238
;; PRIOR FILING DATE: 2000-01-11
;; PRIOR APPLICATION NUMBER: 09/370,265
;; PRIOR FILING DATE: 1999-08-09
;; PRIOR APPLICATION NUMBER: PCT/US99/16180
;; PRIOR FILING DATE: 1999-07-16
;; PRIOR APPLICATION NUMBER: 09/350,836
;; PRIOR FILING DATE: 1999-07-09
;; PRIOR APPLICATION NUMBER: 09/273,447
;; PRIOR FILING DATE: 1999-03-19
;; PRIOR APPLICATION NUMBER: 09/244,444
;; PRIOR FILING DATE: 1999-02-04
;; PRIOR APPLICATION NUMBER: 09/122,449
;; PRIOR FILING DATE: 1998-07-24
;; PRIOR APPLICATION NUMBER: 09/118,205
;; PRIOR FILING DATE: 1998-07-16
;; NUMBER OF SEQ ID NOS: 60
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 25
;; LENGTH: 405
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; US-09-608-285A-25

Query Match 20.5%; Score 49,5; DB 4; Length 405;
Best Local Similarity 40.5%; Pred. No. 21;
Matches 17; Conservative 2; Mismatches 12; Indels 11; Gaps 2;

QY 3 QEAEMAQNAVRLHYGLPVVVVKLLHPPSHWP-----LIKATVGL 40
| | | | | : | | | | : | | | |
Db 99 QGAETVQ-----GLLEVAKDSIPRSHWKTPVVLKATAGL 133

RESULT 14
US-09-608-285A-3
;; Sequence 3, Application US/09608285A
;; Patent No. 6335013
;; GENERAL INFORMATION:
;; APPLICANT: Ford, John
;; APPLICANT: Mulero, Julio
;; APPLICANT: Yeung, George
;; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE

Search completed: July 29, 2002, 16:05:13
Job time: 483 sec

```

; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 28110/36570
; CURRENT APPLICATION NUMBER: US/09/608,285A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/583,231
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-608-285A-3

Query Match 20.5% Score 49.5; DB 4; Length 428;
Best Local Similarity 40.5%; Pred. No. 23;
Matches 17; Conservative 2; Mismatches 12; Indels 11; Gaps 2;

QY 3 QEAEMAQNAVRLHYGLPVPVKLLHPSPHPW----LIRATVGL 40
Db 99 QGAETVQ-----GLLEVAKDIPRSHWKKTPVVLKATAGL 133

RESULT 15
US-09-608-285A-5
; Sequence 5, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 28110/36570
; CURRENT APPLICATION NUMBER: US/09/608,285A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/583,231
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24

```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 16:09:17 ; Search time 70.63 Seconds
(without alignments)
62.581 Million cell updates/sec

Title: US-09-641-104A-12
Perfect score: 241
Sequence: 1 RHQAEAMQNAVRLHYGLPV.....PPSHWPLIKATVGLIRNLAL 46

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

PIR71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	241	100.0	781	2 S35099	beta-catenin - Afr
2	241	100.0	781	2 A38973	beta-catenin - hum
3	241	100.0	781	2 S35091	beta-catenin - mou
4	218	90.5	817	2 S33793	hypothetical prote
5	202	83.8	820	2 S33794	hypothetical prote
6	201	83.4	738	2 S35093	plakoglobin - Afr
7	195	80.9	843	2 T12689	armadillo segment
8	190	78.8	621	2 S35092	plakoglobin - mus
9	190	78.8	744	2 A32905	plakoglobin, desmo
10	188	78.0	806	2 JC4835	beta-catenin - Hyd
11	185	32.6	678	2 T23341	beta-catenin - Cae
12	63.5	26.3	1050	2 T26395	hypothetical prote
13	59.5	24.7	351	2 F90409	GTP binding conser
14	58.5	24.3	578	2 S50446	VAC8 protein - yea
15	56.5	23.4	393	1 E64485	phosphoribosylglyc
16	56.5	23.4	555	2 T40294	hypothetical prote
17	56	23.2	908	2 AE2675	pyruvate, orthophos
18	56	23.2	933	2 C97457	pyruvate, phosphat
19	55.5	23.0	160	1 S26194	plastoquinol--plas
20	55	22.8	224	2 H82658	carboxylesterase X
21	55	22.8	441	2 A84351	hypothetical prote
22	55	22.8	503	2 G81917	probable ubiquinol
23	55	22.8	842	2 T23715	hypothetical prote
24	55	22.8	1160	2 T23713	hypothetical prote
25	55	22.8	1286	2 T23714	hypothetical prote
26	54	22.4	885	2 S66601	pyruvate, orthophos
27	54	22.4	1094	2 T00814	RNA-directed DNA p
28	54	22.4	1124	2 S23820	pol polyprotein -
29	53.5	22.2	396	2 D82394	G0DEF family prote

30 53.5 22.2 899 2 S62428 probable coiled co
31 53 22.0 154 2 S58075 probable olfactory
32 53 22.0 437 2 B82778 cell cycle protein
33 53 22.0 503 2 B81184 ubiquinone biosynt
34 53 22.0 881 2 F72397 pyruvate, orthophos
35 53 22.0 888 2 T31131 pyruvate, orthophos
36 53 22.0 898 2 T37037 pyruvate phosphate
37 53 22.0 900 2 G87431 pyruvate, phosphat
38 53 22.0 930 2 AF3431 pyruvate, orthophos
39 53 22.0 947 2 T02979 pyruvate, orthophos
40 53 22.0 949 1 S55478 pyruvate, orthophos
41 53 22.0 953 2 S12894 pyruvate, orthophos
42 53 22.0 953 2 S56650 pyruvate, orthophos
43 53 22.0 955 2 S56649 pyruvate, orthophos
44 53 22.0 956 1 S53297 pyruvate, orthophos
45 53 22.0 960 2 A71420 pyruvate, orthophos

ALIGNMENTS

RESULT 1

S35099
beta-catenin - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 10-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 21-Jul-2000
C:Accession: S35099
R:McCreary, P.D.; Turck, C.W.; Gumbiner, B.
Science 254, 1359-1361, 1991
A:Title: A homolog of the armadillo protein in Drosophila (plakoglobin) associated wi
A:Reference number: S35099; MUID:92073903
A:Accession: S35099
A:Molecule type: mRNA
A:Residues: 1-781 <MCC>
A:Cross-references: GB:M77013; NID:g214020; PIDN:AAA49670.1; PID:g214021
C:Keywords: cytoskeleton

Query Match 100.0%; Score 241; DB 2; Length 781;
Best Local Similarity 100.0%; Pred. No. 7.5e-24;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RHQAEAMQNAVRLHYGLPVVVVKLLHPPSHWPLIKATVGLIRNLAL 46
Db 474 RHQAEAMQNAVRLHYGLPVVVVKLLHPPSHWPLIKATVGLIRNLAL 519

RESULT 2

A38973
beta-catenin - human
C:Species: Homo sapiens (man)
C:Date: 26-Jan-1996 #sequence_revision 08-Feb-1996 #text_change 08-Oct-1999
C:Accession: A38973; S53356; S31988
R:Huelsken, J.; Birchmeier, W.; Behrens, J.
J. Cell Biol. 127, 2061-2069, 1994
A:Title: E-cadherin and APC compete for the interaction with beta-catenin and the cyt
A:Reference number: A38973; MUID:95105247
A:Accession: A38973
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-781 <HUE>
A:Cross-references: GB:219054; NID:g38519; PIDN:CAA79497.1; PID:g38520
A:Experimental source: placenta
R:Nollet, F.; Berr, G.; Molemans, F.; van Roy, F.
submitted to the EMBL Data Library, June 1995
A:Description: H. sapiens beta-catenin mRNA.
A:Reference number: S55356
A:Accession: S55356
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-781 <NOL>
A:Cross-references: EMBL:X87838; NID:g1154853; PIDN:CAA61107.1; PID:g860988
C:Comment: The cytoplasmic protein beta-catenin normally binds to a region in the cyt

microfilament network.

C:Comment: Cellular levels of beta-catenin are regulated in part by the adenomatous polyd cancerous cell growth.

C:Genetics:

A:Gene: GDB:CTNNB1; CTNNB

A:Cross-references: GDB:141922; OMIM:116806

A:Map position: 3p22-3p21.3

C:Keywords: apoptosis; carcinogenesis; cell adhesion; cytosol

F:151-676/Region: 40-residue repeats

Query Match 100.0%; Score 241; DB 2; Length 781;
Best Local Similarity 100.0%; Pred. No. 7.5e-24;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHQEAEMAGNAVRLHYGLPVVVKLLHPPSHWPLIKATVGLIRNAL 46
|||||
DB 474 RHQEAEMAGNAVRLHYGLPVVVKLLHPPSHWPLIKATVGLIRNAL 519

RESULT 3

S35091 beta-catenin - mouse

C:Species: Mus musculus (house mouse)

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999

C:Accession: S35091

R:Butz, S.; Stappert, J.; Weissig, H.; Kemler, R.

Science 257, 1142-1144, 1992

A:Title: Plakoglobin and beta-catenin: distinct but closely related.

A:Reference number: S35091; MUID:92376536

A:Accession: S35091

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-781 <BUT>

A:Cross-references: EMBL:M90364; NID:gl192141; PIDN:AAA37280.1; PID:gl192142

C:Keywords: cytoskeleton

Query Match 100.0%; Score 241; DB 2; Length 781;
Best Local Similarity 100.0%; Pred. No. 7.5e-24;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RHQEAEMAGNAVRLHYGLPVVVKLLHPPSHWPLIKATVGLIRNAL 46
|||||
DB 474 RHQEAEMAGNAVRLHYGLPVVVKLLHPPSHWPLIKATVGLIRNAL 519

RESULT 4

S33793

hypotheical protein - spoonworm (Urechis caupo)

C:Species: Urechis caupo

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Aug-1997

C:Accession: S33793

R:Rosenthal, E.

Biochim. Biophys. Acta 1173, 337-341, 1993

A:Title: Identification of homologues to beta-catenin/plakoglobin/armadillo in two inver

A:Reference number: S33793; MUID:93305730

A:Accession: S33793

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-817 <ROS>

Query Match 90.5%; Score 218; DB 2; Length 817;
Best Local Similarity 91.3%; Pred. No. 9e-21;
Matches 42; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RHQEAEMAGNAVRLHYGLPVVVKLLHPPSHWPLIKATVGLIRNAL 46
|||||
DB 496 RHQEAEMAGNAVRLHYGLPVVVKLLHPPSRWPLIKAVVGLIRNAL 541

RESULT 5

S33794

hypotheical protein - sea urchin (Tripneustes gratilla)

C:Species: Tripneustes gratilla

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Aug-1997

C:Accession: S33794

R:Rosenthal, E.

Biochim. Biophys. Acta 1173, 337-341, 1993

A:Title: Identification of homologues to beta-catenin/plakoglobin/armadillo in two in

A:Reference number: S33793; MUID:93305730

A:Accession: S33794

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-820 <ROS>

Query Match 83.8%; Score 202; DB 2; Length 820;
Best Local Similarity 82.6%; Pred. No. 1.2e-18;
Matches 38; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 RHQEAEMAGNAVRLHYGLPVVVKLLHPPSHWPLIKATVGLIRNAL 46
|||||
DB 490 RHQEAEMAGNAVRLHYGLPVVVKLLHPPSRWPLIKATVGLIRNAL 535

RESULT 6

S35093 plakoglobin - African clawed frog

C:Species: Xenopus laevis (African clawed frog)

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jul-2000

C:Accession: S35093; S24636

R:Fouquet, B.; Zimbelmann, R.; Franke, W.W.

Differentiation 51, 187-194, 1992

A:Title: Identification of plakoglobin in oocytes and early embryos of Xenopus laevis

A:Reference number: S35093; MUID:93093332

A:Accession: S35093

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-738 <FOU>

A:Cross-references: EMBL:M95593; NID:g214656; PIDN:AAA49931.1; PID:g214657

R:Demarais, A.A.

submitted to the EMBL Data Library, June 1992

A:Reference number: S24636

A:Accession: S24636

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 133-184, 'V', 186-225, 'T', 227-292 <DEM>

A:Cross-references: EMBL:X67078; NID:g65252; PID:g65253

C:Keywords: cytoskeleton

Query Match 83.4%; Score 201; DB 2; Length 738;
Best Local Similarity 76.1%; Pred. No. 1.4e-18;
Matches 35; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 RHQEAEMAGNAVRLHYGLPVVVKLLHPPSHWPLIKATVGLIRNAL 46
|||||
DB 463 RHQEAEMAGNAVRLHYGLPVVVKLLHPPSQWPLIKATVGLIRNAL 506

RESULT 7

T12689 armadillo segment polarity protein - fruit fly (Drosophila melanogaster)

A:Alternate names: protein 86E4.6

C:Species: Drosophila melanogaster

C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 21-Jul-2000

C:Accession: T12689; A31861

R:Ferraz, C.; Vidal, S.; Brun, C.; Bucheton, A.; Demaille, J.G.

submitted to the EMBL Data Library, January 1998

A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.

A:Reference number: Z17572

A:Accession: T12689

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

```
A:Map position: 7pter-7qter
C:Keywords: cytoskeleton

Query Match      78.8%   Score 190;   DB 2;   Length 744;
Best Local Similarity 73.9%;   Pred. No. 4.2e-17;
Matches 34; Conservative 4; Indels

QY          1 RHQAEMAQNRAVRHYGLPVVVVKLLHPSPHWPLIKATVGLIRNAL 46
            | | | | | | | | | | | | | | | | | | | | | |
Db         464 RHPEAEMAQNSRVLYNGYIPAVKLLNQNPWPLVKATIGLIRNAL 509

RESULT 10
JC4835
beta-catenin - Hydra magnipapillata
c:Species: Hydra magnipapillata
```

C:Accession: JC4835
R:Hobmayer, E.; Hatta, M.; Fischer, R.; Fujisawa, T.; Holstein, T.
Gene 172, 155-159, 1996
A:Title: Identification of a Hydra homologue of the beta-catenin.
A:Reference number: JC4835; MUID:96257271
A:Accession: JC4835
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA

```

Query Match          78.0% ; Score 188; DB 2; Length 806;
Best Local Similarity 73.9%; Pred. No. 8.5e-17;
Matches    34; Conservative   6; Mismatches    6; Indels

QY      1 RHQAENMAQNAAVRLLHYGLPVVVVKKLHPPSPHWPLIKATVGLIRNLAL 46
      :|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db       533 RHPDAEHAENGVRLLHYGPIILVKLLNPPSRWPLIKAVVGLIRNLGL 578

RESULT      11;
              723341
              T

```

C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change: 0
C:Accession: T23341; T42221
R:Harris, B.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19729
A:Accession: T23341
A:Status: preliminary; translated from GB/EMBL/DBJ

A; Residues: 1-678 <WIL>
A; Cross-references: EMBL: Z81564; PIDN: CAB04572.1; GSPDB: GN000119;
A; Status: preliminary; translated from GB/EMBL/DBJ
A; Molecule type: mRNA
A; Residues: 1-678 <COS>
A; Cross-references: EMBL: AF016853; PIDN: AAB94552.1
C; Genetics:
A; Gene: hmp-2; CESP: K0504.6
A; Map position: 1
A; Introns: 14/1; 235/1; 290/1; 620/3
Query Match
32.6% Score 78.5 DB 2 Length 678

```

Best Local Similarity 43.5%;   Pred. No. 0.024;
Matches 20; Conservative 7; Mismatches 18; Indels 1; Gaps 1;

QY      1 RHQEAEMQNAVRLHYGLPVVVVKLLHPPSHWPLIKATVGLIRNAL 46
      || || || : || :: : || :: || :: || :: || :: ||
Ddb     404 RHSLAEAAQSELRFCQAFVILDQLE-TLRTPVKAALGVIRNSAL 448

RESULT 12
T26395
hypothetical protein Y105C5B.s - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T26395
R:McMurray, A.
submitted to the EMBL Data Library, September 1999
A:Reference number: Z20208
A:Accession: T26395
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-1050 <WILL>
A:Cross-references: EMBL:AL110479; NID:el542153; PIDN:CAB54367.1; CESP:Y105C5B.s
A:Experimental source: clone Y105C5B
C:Genetics:
A:Gene: CESP:Y105C5B.s
A:Introns: 40/3; 105/1; 217/3; 278/1; 422/1; 482/1; 850/1; 903/1; 959/2

```

Query Match 26.3%; Score 63.5; DB 2; Length 1050;
Best Local Similarity 41.7%; Pred. No. 3.8;
Matches 15; Conservative 9; Mismatches 11; Indels 1; Gaps 1;

QY 11 AVRLHGLPVPVVKLLHPPSHWPLIKATVGLIRNLAL 46
|||: ||||:|:|: | : :: | : ||||:
DB 807 AVRVEKGLPVLVLIPLPEDF-VYCAVATALRNLA 841

RESULT 13

F90409
GTP binding conserved hypothetical protein SS02385 [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C:Accession: F90409
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, E.
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: F90409
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-351 <KUR>
A:A:Cross-references: GB:AE006641; NID:g13815687; PIDN:AAK42533.1; GSPDB:GN00155
C:Genetics:
A:Gene: SS02385

Query Match	24.7%	Score 59.5	DB 2	Length 351
Best Local Similarity	54.2%	Pred. No. 3.8		
Matches 13	Conservative 5	Mismatches 5	Indels 1	Gaps 1
QY	20	VVVKLLHPSPHWPLIKATVGLI	43	
DB	120	VOTLVNPPKSLPLSK-TIGLI	142	

RESULT 14
S50446
VAC8 protein - yeast (Saccharomyces cerevisiae)
Alternate names: protein YEL013w
C:Species: Saccharomyces cerevisiae
C.Date: 28-Nov-1993 #sequence revision 24-Feb-1995 #text change 29-Oct-1999

C.Accession: S50446;
R.Dietrich, F.S.
submitted to the EMBL Data Library, December 1994
A:Description: Saccharomyces cerevisiae chromosome V cosmids 9871, 8199, 9867, 9495 a
A:Reference number: S50428
A:Accession: S50446
A:Molecule type: DNA
A:Residues: 1-578 <DIE>
A:Cross-references: EMBL:U18530; NID:g602367; PID:g602380; GSPDB:GN00005; MIPS:YEL013
C:Genetics:
A:Gene: SGD:VAC8; MIPS:YEL013w
A:Cross-references: SGD:S0000739; MIPS:YEL013w
A:Map position: 5L
C:Function:
A:Description: required for vacuole inheritance and protein targeting from the cytopl
C:Keywords: yeast vacuole

Query Match 24.3%; Score 58.5; DB 2; Length 578;
Best Local Similarity 46.7%; Pred. NO. 9;
Matches 14; Conservative 7; Mismatches 8; Indels 1; Gaps 1;

QY 17 GLPFWVKLLHPPSHWPLIKATVGLIRNLAL 46
111 1111 1 111 1 11111
Db 203 GLPFWVKLLHPPSHWPLIKATVGLIRNLAL 321

RESULT 15
E64485 .
phosphoribosylglycinamide formyltransferase 2 (EC 2.1.2.-) MJ1486 [similarity] - Meth
C.Species: Methanococcus jannaschii
C.Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C.Accession: E64485
R;Baik, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak
/ Rulif, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek,
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese
A;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannasc
A;Reference number: A64300; MUID:96337999
A;Accession: E64485
A;Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-393 <BUIL>
A;Cross-references: GB:U67589; GB:L77117; NID:g2826417; PIDN:AAB99496.1; PID:g1592123
C;Genetics:
A;Map position: REV1458908-1457727
A;Start codon: TTG
A;Function:

	Query Match	23.4%	Score 56.5;	DB 1;	Length 393;
	Best Local Similarity	45.9%	Pred. No. 11;		
	Matches 17;	Conservative	4;	Mismatches 13;	
				Indels 3;	Gaps 2;
Qy	3	QEAEMAQNAVRLHYGLPVPVKLLHP--PSHMPLIKATV	38		
qh	298	QEMFEFRTHVRLGLPVPVKLLIHPGASH--VIAKEI	332		

Search completed: July 29, 2002, 16:09:18
Job time: 728 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 16:09:56 ; Search time 34.24 Seconds
(without alignments)
52.018 Million cell updates/sec

Title: US-09-641-104a-12

Perfect score: 241

Sequence: 1 RHQEAQAQNAVRLHYGLPV.....PPSHWPLKATVGLIRNLAL 46

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	241	100.0	781	1	CTNB_HUMAN
2	241	100.0	781	1	CTNB_MOUSE
3	241	100.0	781	1	CTNB_RAT
4	241	100.0	781	1	CTNB_XENLA
5	218	90.5	818	1	CTNB_XENLA
6	202	83.8	820	1	CTNB_TRIGR
7	201	83.4	738	1	PLAK_XENLA
8	198	82.2	813	1	ARM_MUSDO
9	195	80.9	843	1	ARM_MOUSE
10	190	78.8	621	1	PLAK_MOUSE
11	190	78.8	743	1	PLAK_HUMAN
12	58.5	24.3	578	1	VAC8_YEAST
13	56.5	23.4	313	1	B3GT_DROME
14	56.5	23.4	393	1	PURT_METJA
15	55.5	23.0	160	1	PETD_SYNP2
16	55	22.8	503	1	UBIB_NEIMA
17	55	22.8	749	1	PEX_MOUSE
18	55	22.8	898	1	PODK_RHIME
19	54	22.4	885	1	PODK_ENTHI
20	54	22.4	1124	1	POL_FIVSD
21	53.5	22.2	899	1	YABD_SCHPO
22	53.5	22.2	962	1	ARVC_HUMAN
23	53.5	22.2	969	1	ARVC_MOUSE
24	53	22.0	503	1	UBIB_NEIMB
25	53	22.0	949	1	PODK_MESCR
26	53	22.0	953	1	PODK_FLABI
27	53	22.0	953	1	PODK_FLATR
28	53	22.0	955	1	PODK_FLATR
29	53	22.0	956	1	PODK_FLAPR
30	52	21.6	749	1	PEX_HUMAN
31	51.5	21.4	1124	1	POL_FIVT2
32	51	21.2	191	1	Y023_BUCAI
33	51	21.2	308	1	YFER_ECOLI

RESULT 1

CTNB_HUMAN

ID CTNB_HUMAN STANDARD; PRT; 781 AA.
AC P35222;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Beta-catenin.
GN CTNNB1 OR CTNNB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Placenta;
RC MEDLINE=95105247; PubMed=7806582;
RA Huelken J., Birchmeier W., Behrens J.;
RT "E-cadherin and APC compete for the interaction with beta-catenin and the cytoskeleton.";
RL J. Cell Biol. 127:2061-2069(1994).
[2]
RP REVIEW.
RX MEDLINE=20145417; PubMed=10679188;
RA Kikuchi A.;
RL "Regulation of beta-catenin signaling in the Wnt pathway.";
[3]
RX X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 133-664.
RX MEDLINE=20578806; PubMed=11136974;
RA Graham T.A., Weaver C., Mao F., Kimelman D., Xu W.;
RL "Crystal structure of a beta-catenin/Tcf complex.";
Cell 103:885-896(2000).
[1]- FUNCTION: INVOLVED IN THE REGULATION OF CELL ADHESION AND IN SIGNAL TRANSDUCTION THROUGH THE WNT PATHWAY.
[1]- SUBUNIT: THE CYTOSOLIC FORM BOUND TO AXIN ALLOWS THE STIMULATION OF ITS DEGRADATION. THE CYTOSOLIC FORM ALSO INTERACTS WITH APC AND E-CADHERIN. THE NUCLEAR FORM BINDS TCF/LEF-1 AND MAY ALSO BIND PONTIN52 AND DUPLIN. CAN HETERODIMERIZE WITH ALPHA-CATENIN AND ALSO FORM A TRIMERIC COMPLEX WITH CADHERIN. MEDIATES THE ANCHORAGE OF THE CADHERIN COMPLEX TO THE CORTICAL ACTIN CYTOSKELETON.
[1]- SUBCELLULAR LOCATION: CYTOPLASMIC WHEN IT IS UNSTABILIZED (HIGH LEVEL OF PHOSPHORYLATION). TRANSLOCATES TO THE NUCLEUS WHEN IT IS STABILIZED (LOW LEVEL OF PHOSPHORYLATION).
[1]- PTM: PROBABLY PHOSPHORYLATED BY GSK-3B. WNT-DEPENDENT ACTIVATION OF DVL ANTAGONIZES THE ACTION OF GSK-3B. PHOSPHORYLATION OF BETA-CATENIN IS REDUCED AND THE PROTEIN IS NO LONGER DEGRADED, RESULTING IN ITS ACCUMULATION IN CYTOPLASM.
[1]- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
[1]- SIMILARITY: CONTAINS 12 ARM REPEATS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

34 51 21.2 407 1 HEMO_RHOSH
35 51 21.2 649 1 VE1_HPV6A
36 51 21.2 649 1 VE1_HPV6B
37 50.5 21.0 261 1 COX3_PELSU
38 50.5 21.0 594 1 YC9F_SCHPO
39 50.5 21.0 694 1 TKTL_RHIME
40 50 20.7 451 1 VPS9_YEAST
41 50 20.7 498 1 CIKW_DROME
42 50 20.7 552 1 UBIB_XYLFA
43 50 20.7 1849 1 T172_HUMAN
44 49.5 20.5 126 1 C11X_HUMAN
45 49.5 20.5 428 1 ENP5_HUMAN

ALIGNMENTS

Q06965 rhodobacter
Q84293 human papil
P03113 human papil
O79676 pelomedusa
Q09889 schizosacch
P38333 rhizobium m
P54787 saccharomyc
P17972 drosophila
Q9pce8 xylella fas
O14981 homo sapien
Q15332 homo sapien
O75356 homo sapien

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: X87838; CAA61107.1; -;
 CC EMBL: Z19054; CAA79497.1; -;
 DR PIR: S31988; S31988.
 DR PDB: 1G3J; 17-JAN-01.
 DR TRANSFAC: T02872; -;
 DR MIM: 116806; -;
 DR InterPro: IPR000225; Armadillo.
 DR Pfam: PF00514; Armadillo_seg; 12.
 DR SMART: SM00185; ARM; 11.
 DR PROSITE: PS50176; ARM_REPEAT; 9.
 KW Cell adhesion; Cytoskeleton; Structural protein; Nuclear protein;
 FT Repeat; 3D-structure.
 FT REPEAT 151 191
 FT REPEAT 193 234
 FT REPEAT 235 276
 FT REPEAT 277 318
 FT REPEAT 319 360
 FT REPEAT 361 389
 FT REPEAT 400 441
 FT REPEAT 442 484
 FT REPEAT 489 530
 FT REPEAT 531 571
 FT REPEAT 594 636
 FT REPEAT 637 666
 FT SEQUENCE 781 AA; 85496 MW; CB78F165A3EEF86E CRC64;

Query Match 100.0%; Score 241; DB 1; Length 781;

Best Local Similarity 100.0%; Pred. No. 2.7e-24; Indels 0; Gaps 0;
 Matches 46; Conservative 0; Mismatches 0;

QY 1 RHQEAQNAQNAVRHLHYGLPVVVKLLHPPSHWPLIKATVGLIRNAL 46
 |||||
 DB 474 RHQEAQNAQNAVRHLHYGLPVVVKLLHPPSHWPLIKATVGLIRNAL 519

RESULT 2

CTNB_MOUSE ID CTNB_MOUSE STANDARD; PRT; 781 AA.
 AC Q02248;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Beta-catenin.
 GN CTNNB1 OR CATNB.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxId=10090;
 FN [1]
 SE SEQUENCE FROM N.A.
 RX MEDLINE=92376536; PubMed=1509266;
 RA Butz S., Stappert J., Weissig H., Kemler R.;
 RT "Plakoglobin and beta-catenin: distinct but closely related.";
 RL Science 257:1142-1144(1992).
 RN [2]
 RN REVIEW.
 RP MEDLINE=20145417; PubMed=10679188;
 RX Kikuchi A.;
 RA "Regulation of beta-catenin signaling in the Wnt pathway.";
 RT Biochem. Biophys. Res. Commun. 268:243-248(2000).
 RL [3]
 RN X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 150-665.
 RP MEDLINE=97442350; PubMed=9298899;
 RX Huber A.H., Nelson W.J., Weiss W.I.;
 RA "Three-dimensional structure of the armadillo repeat region of beta-catenin.";
 RT Cell 90:871-882(1997).
 RL [4]
 RN X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 118-149 IN COMPLEX WITH

RP CTNNB1
 RX MEDLINE=20337986; PubMed=10882138;
 RA Pokutta S., Weis W.I.;
 RT "Structure of the dimerization and beta-catenin-binding region of alpha-catenin.";
 RL Mol. Cell 5:533-543(2000).
 CC - FUNCTION: INVOLVED IN THE REGULATION OF CELL ADHESION AND IN SIGNAL TRANSDUCTION THROUGH THE WNT PATHWAY.
 CC - SUBUNIT: THE CYTOSOLIC FORM BOUND TO AXIN ALLOWS THE STIMULATION OF ITS DEGRADATION. THE CYTOSOLIC FORM ALSO INTERACTS WITH APC AND E-CADHERIN. THE NUCLEAR FORM BINDS TCF/LEF-1 AND MAY ALSO BIND POINTIN52 AND DUPLIN. CAN HETERODIMERIZE WITH ALPHA-CATENIN AND ALSO FORM A TRIMERIC COMPLEX WITH CADHERIN. MEDIATES THE ANCHORAGE OF THE CADHERIN COMPLEX TO THE CORTICAL ACTIN CYTOSKELETON.
 CC - SUBCELLULAR LOCATION: CYTOPLASMIC WHEN IT IS UNSTABILIZED (HIGH LEVEL OF PHOSPHORYLATION). TRANSLOCATES TO THE NUCLEUS WHEN IT IS STABILIZED (LOW LEVEL OF PHOSPHORYLATION).
 CC - PTM: PROBABLY PHOSPHORYLATED BY GSK-3B. WNT-DEPENDENT ACTIVATION OF DVL ANTAGONIZES THE ACTION OF GSK-3B. PHOSPHORYLATION OF BETA-CATENIN IS REDUCED AND THE PROTEIN IS NO LONGER DEGRADED.
 CC - RESULTING IN ITS ACCUMULATION IN CYTOPLASM.
 CC - SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
 CC - SIMILARITY: CONTAINS 12 ARM REPEATS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use, by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: M90364; AAA37280.1; -;
 DR PIR: S35091; S35091.
 DR PDB: 2BCT; 15-OCT-97.
 DR PDB: 3BCT; 19-NOV-97.
 DR PDB: 1DOW; 12-JUL-00.
 DR TRANSFAC: T02984; -;
 DR MGD: MGI:88276; Catnb.
 DR InterPro: IPR000225; Armadillo.
 DR Pfam: PF00514; Armadillo_seg; 12.
 DR SMART: SM00185; ARM; 11.
 DR PROSITE: PS50176; ARM_REPEAT; 9.
 KW Cell adhesion; Cytoskeleton; Structural protein; Nuclear protein;
 FT Repeat; 3D-structure.
 FT REPEAT 151 191
 FT REPEAT 193 234
 FT REPEAT 235 276
 FT REPEAT 277 318
 FT REPEAT 319 360
 FT REPEAT 361 389
 FT REPEAT 400 441
 FT REPEAT 442 484
 FT REPEAT 489 530
 FT REPEAT 531 571
 FT REPEAT 594 636
 FT REPEAT 637 666
 FT SEQUENCE 781 AA; 85470 MW; D708F170A3FBED6E CRC64;

Query Match 100.0%; Score 241; DB 1; Length 781;

Best Local Similarity 100.0%; Pred. No. 2.7e-24; Indels 0; Gaps 0;
 Matches 46; Conservative 0; Mismatches 0;

QY 1 RHQEAQNAQNAVRHLHYGLPVVVKLLHPPSHWPLIKATVGLIRNAL 46
 |||||
 DB 474 RHQEAQNAQNAVRHLHYGLPVVVKLLHPPSHWPLIKATVGLIRNAL 519

RESULT 3

CTNB_MOUSE ID CTNB_MOUSE STANDARD; PRT; 781 AA.
 AC Q09082;

DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Beta-catenin.
GN CTNNB1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE=Testis;
RX MEDLINE=94428593; PubMed=10497305;
RA Chung S.S.W., Lee W.M., Cheng C.Y.;
RT "Study on the formation of specialized inter-Sertoli cell junctions in
RT vitro.";
RL J. Cell. Physiol. 181:258-272(1999).
CC -1- FUNCTION: INVOLVED IN THE REGULATION OF CELL ADHESION AND IN
CC SIGNAL TRANSDUCTION THROUGH THE WNT PATHWAY (BY SIMILARITY).
CC OF ITS DEGRADATION. THE CYTOSOLIC FORM ALSO INTERACTS WITH APC AND
CC E-CADHERIN. THE NUCLEAR FORM BINDS TCF/LEF-1 AND MAY ALSO BIND
CC PONTIN52 AND DUPLIN. CAN HETERODIMERIZE WITH ALPHA-CATENIN AND
CC ALSO FORM A TRIMERIC COMPLEX WITH CADHERIN. MEDIATES THE ANCHORAGE
CC OF THE CADHERIN COMPLEX TO THE CORTICAL ACTIN CYTOSKELETON (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC WHEN IT IS UNSTABILIZED (HIGH
CC LEVEL OF PHOSPHORYLATION). TRANSLOCATES TO THE NUCLEUS WHEN IT IS
CC STABILIZED (LOW LEVEL OF PHOSPHORYLATION) (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE TESTIS.
CC -1- DEVELOPMENTAL STAGE: HIGHLY EXPRESSED AT E30-60 DAY DPC IN THE
CC TESTIS. REDUCED EXPRESSION AT E90 DAY DPC.
CC -1- PTM: PROBABLY PHOSPHORYLATED BY GSK-3B. WNT-DEPENDENT ACTIVATION
CC OF DVL ANTAGONIZES THE ACTION OF GSK-3B. PHOSPHORYLATION OF BETA-
CC CATENIN IS REDUCED AND THE PROTEIN IS NO LONGER DEGRADED,
CC RESULTING IN ITS ACCUMULATION IN CYTOPLASM (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -1- SIMILARITY: CONTAINS 12 ARM REPEATS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AF121265; AAD28504.1; -
DR HSSP; Q02248; 2BCT.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 12.
DR SMART; SM00185; ARM; 11.
DR PROSITE; PS50176; ARM_REPEAT; 9.
KW Cell adhesion; Cytoskeleton; Structural protein; Nuclear protein;
KW Repeat.
FT REPEAT 151 191 ARM 1.
FT REPEAT 193 234 ARM 2.
FT REPEAT 235 276 ARM 3.
FT REPEAT 277 318 ARM 4.
FT REPEAT 319 360 ARM 5.
FT REPEAT 361 389 ARM 6.
FT REPEAT 400 441 ARM 7.
FT REPEAT 442 484 ARM 8.
FT REPEAT 489 530 ARM 9.
FT REPEAT 531 571 ARM 10.
FT REPEAT 594 636 ARM 11.
FT REPEAT 637 666 ARM 12.
SQ SEQUENCE 781 AA; 85454 MW; 9C29186B5DBD54B87 CRC64;

Query Match 100.0%; Score 241; DB 1; Length 781;
Best Local Similarity 100.0%; Pred. No. 2.7e-24;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RHOEAEQAQNAVRLHYGLPVVVKLLHPPSHWPLIKATVGLIRNLAL 46
3 |||||
DB 474 RHOEAEQAQNAVRLHYGLPVVVKLLHPPSHWPLIKATVGLIRNLAL 519
|||||
RESULT 4
CTNB_XENLA STANDARD; PRT; 781 AA.
ID CTNB_XENLA
AC P26233;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Beta-catenin.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.; AND PARTIAL SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=92073903; PubMed=1962194;
RA McCrea P.D., Turk C.W., Gumbiner B.M.;
RT "A homolog of the armadillo protein in Drosophila (plakoglobin)
RT associated with E-cadherin.";
RL Science 254:1359-1361(1991).
CC -1- FUNCTION: BINDS TO THE CYTOPLASMIC DOMAIN OF THE CELL-CELL
CC ADHESION MOLECULE E-CADHERIN, AND PERHAPS TO OTHER (MEMBRANE)
CC PROTEINS. THE ASSOCIATION OF CATENINS TO CADHERINS PRODUCES A
CC COMPLEX WHICH IS LINKED TO THE ACTIN FILAMENT NETWORK, AND WHICH
CC SEEMS TO BE OF PRIMARY IMPORTANCE FOR CADHERINS CELL-ADHESION
CC PROPERTIES: ESSENTIAL COMPONENT OF THE VERTEBRATE WNT-SIGNALING
CC PATHWAY AND IS REQUIRED FOR FORMATION OF THE EMBRYONIC AXIS.
CC -1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -1- SIMILARITY: CONTAINS 10 ARM REPEATS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; W77013; AAA49670.1; -
DR PIR; S35099; S35099.
DR HSSP; Q02248; 2BCT.
DR TRANSFAC; T03026; -
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 12.
DR SMART; SM00185; ARM; 11.
DR PROSITE; PS50176; ARM_REPEAT; 9.
KW Cell adhesion; Cytoskeleton; Structural protein; Repeat.
KW Repeat.
FT REPEAT 141 180 ARM 1.
FT REPEAT 181 225 ARM 2.
FT REPEAT 225 264 ARM 3.
FT REPEAT 267 306 ARM 4.
FT REPEAT 351 390 ARM 5.
FT REPEAT 391 429 ARM 6.
FT REPEAT 432 473 ARM 7.
FT REPEAT 479 519 ARM 8.
FT REPEAT 521 562 ARM 9.
FT REPEAT 584 623 ARM 10.
FT REPEAT 625 664 ARM 11.
SQ SEQUENCE 781 AA; 85449 MW; 3ECD27232239F799 CRC64;

Query Match 100.0%; Score 241; DB 1; Length 781;
Best Local Similarity 100.0%; Pred. No. 2.7e-24;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 RHOEAEQAQNAVRLHYGLPVVVKLLHPPSHWPLIKATVGLIRNLAL 46
|||||

Db 474 RHQEAEMAQNAVRLHYGLPVVVKLLHPPSHWPLIKATVGLIRNLAL 519

```

RESULT 5
CTNB_URECA
ID CTNB_URECA STANDARD; PRT; 818 AA.
AC P35224;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Beta-catenin.
OS Urechis caupo (Innkeeper worm) (Spoonworm).
OC Eukaryota; Metazoa; Echiura; Xenopneusta; Urechidae; Urechis.
OX NCBI_TaxID=6431;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93305730; PubMed=8318544;
RA Rosenthal E.T.;
RT "Identification of homologues to beta-catenin/plakoglobin/armadillo
in two invertebrates, Urechis caupo and Tripneustes gratilla.";
RL Biochim. Biophys. Acta 1173:337-341(1993).
CC -1- FUNCTION: BINDS TO THE CYTOPLASMIC DOMAIN OF THE CELL-CELL
ADHESION MOLECULE E-CADHERIN, AND PERHAPS TO OTHER (MEMBRANE)
PROTEINS. THE ASSOCIATION OF CATENINS TO CADHERINS PRODUCES A
COMPLEX WHICH IS LINKED TO THE ACTIN FILAMENT NETWORK, AND WHICH
SEEMS TO BE OF PRIMARY IMPORTANCE FOR CADHERINS CELL-ADHESION
PROPERTIES.
CC -1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -1- SIMILARITY: CONTAINS 7 ARM REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; L10355; AAA30330.1; -.
DR HSP; 002248; 28CT.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 12.
DR SMART; SM00185; ARM; 11.
DR PROSITE; PS00176; ARM_REPEAT; 9.
KW Cell adhesion; Cytoskeleton; Structural protein; Repeat.
FT REPEAT 164 203 ARM 1.
FT REPEAT 248 287 ARM 2.
FT REPEAT 412 451 ARM 3.
FT REPEAT 454 495 ARM 4.
FT REPEAT 501 541 ARM 5.
FT REPEAT 543 582 ARM 6.
FT REPEAT 648 687 ARM 7.
SQ SEQUENCE 818 AA; 89070 MW; 1DF174BEEF745C1D CRC64;

```

Query Match 90.5%; Score 218; DB 1; Length 818;
 Best Local Similarity 91.3%; Pred. No. 3.2e-21;
 Matches 42; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RHQEAEMAQNAVRLHYGLPVVVKLLHPPSHWPLIKATVGLIRNLAL 46
 || |||||
 Db 496 RHGAEAMAQNAVRLHYGLPVVVKLLHPPSRWPLIKAVVGLIRNLAL 541

```

RESULT 6
CTNB_TRIGR
ID CTNB_TRIGR STANDARD; PRT; 820 AA.
AC P35223;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Beta-catenin.
OS Tripneustes gratilla (Hawaiian sea urchin).

```

```

OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Temnopleuroidea; Toxopneustidae;
OC Tripneustes.
OX NCBI_TaxID=7673;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93305730; PubMed=8318544;
RA Rosenthal E.T.;
RT "Identification of homologues to beta-catenin/plakoglobin/armadillo
in two invertebrates, Urechis caupo and Tripneustes gratilla.";
RL Biochim. Biophys. Acta 1173:337-341(1993).
CC -1- FUNCTION: BINDS TO THE CYTOPLASMIC DOMAIN OF THE CELL-CELL
ADHESION MOLECULE E-CADHERIN, AND PERHAPS TO OTHER (MEMBRANE)
PROTEINS. THE ASSOCIATION OF CATENINS TO CADHERINS PRODUCES A
COMPLEX WHICH IS LINKED TO THE ACTIN FILAMENT NETWORK, AND WHICH
SEEMS TO BE OF PRIMARY IMPORTANCE FOR CADHERINS CELL-ADHESION
PROPERTIES.
CC -1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -1- SIMILARITY: CONTAINS 10 ARM REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; L10354; AAA30089.1; -.
DR PIR; S33794; S33794.
DR HSP; 002248; IDOM.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 12.
DR SMART; SM00185; ARM; 12.
DR PROSITE; PS00176; ARM_REPEAT; 9.
KW Cell adhesion; Cytoskeleton; Structural protein; Repeat.
FT REPEAT 157 196 ARM 1.
FT REPEAT 199 239 ARM 2.
FT REPEAT 241 280 ARM 3.
FT REPEAT 283 322 ARM 4.
FT REPEAT 367 405 ARM 5.
FT REPEAT 406 445 ARM 6.
FT REPEAT 448 489 ARM 7.
FT REPEAT 495 535 ARM 8.
FT REPEAT 603 642 ARM 9.
FT REPEAT 644 683 ARM 10.
SQ SEQUENCE 820 AA; 89361 MW; 57255E0F57795FD3 CRC64;

```

Query Match 83.8%; Score 202; DB 1; Length 820;
 Best Local Similarity 82.6%; Pred. No. 4.2e-19;
 Matches 38; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 RHQEAEMAQNAVRLHYGLPVVVKLLHPPSHWPLIKATVGLIRNLAL 46
 || |||||
 Db 490 RHGAEAMGONTVRLNYGIPVVKLLHPPSRWPLIKATVGLIRNLAL 535

```

RESULT 7
PLAK_XENLA
ID PLAK_XENLA STANDARD; PRT; 738 AA.
AC P30998;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Junction plakoglobin (Desmoplakin III).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.

```


DE Armadillo segment polarity protein.
GN ARM OR EG:86E4.6 OR CGI1579.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A. (CYTOPLASMIC ISOFORM).
RC STRAIN=OREGON-R;
RX MEDLINE=89211895; PubMed=2707602;
RA Riggelman B., Wieschaus E., Schedl P.;
RT "Molecular analysis of the armadillo locus: uniformly distributed
RT transcripts and a protein with novel internal repeats are associated
RT with a Drosophila segment polarity gene.";
RL Genes Dev. 3:96-113(1989).
RN [2]
RP SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, AND ALTERNATIVE
RP SPLICING.
RC TISSUE=Head;
RX MEDLINE=98298928; PubMed=9635189;
RA Loureiro J., Peifer M.;
RT "Roles of Armadillo, a Drosophila catenin, during central nervous
RT system development.";
RL Curr. Biol. 8:622-632(1998).
RN [3]
RP SEQUENCE FROM N.A. (CYTOPLASMIC ISOFORM).
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Abayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Flosser C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleeb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [4]
RP SEQUENCE FROM N.A. (CYTOPLASMIC ISOFORM).
RC STRAIN=OREGON-R;
RX MEDLINE=20196011; PubMed=10731137;
RA Benos P.V., Gatt M.K., Ashburner M., Murphy L., Harris D., Cadieu E.,
RA Bartell B.G., Ferraz C., Vidal S., Brun C., Demailles J.,

DRano S., Gloux S., Lelaure V., Mottier S., Galibert F., Borkova D.,
RA Minana B., Kafatos F.C., Louis C., Siden-Kiamos I., Bolshakov S.,
RA Papagiannakis G., Spanos L., Cox S., Madueno E., de Pablos B.,
RA Modolell J., Peter A., Schoettler P., Werner M., Mourikoti F.,
RA Beilert N., Dowe G., Schaefer U., Jaecle H., Bucheton A.,
RA Callister D.M., Campbell L.A., Darlamitsou A., Henderson N.S.,
RA McMillan P.J., Salles C., Tait E.A., Valentini P., Saunders R.D.C.,
RA Glover D.M.;
RT "From sequence to chromosome: the tip of the X chromosome of D.
RT melanogaster.";
RL Science 287:2220-2222(2000).
RN [5]
RP PHOSPHORYLATION.
RX MEDLINE=95113174; PubMed=7529201;
RA Peifer M., Pai L.-M., Casey M.;
RT "Phosphorylation of the Drosophila adherens junction protein
RT Armadillo: roles for wingless signal and zeste-white 3 kinase.";
RL Dev. Biol. 166:543-556(1994).
RN [6]
RP FUNCTION: NEURAL ISOFORM MAY ASSOCIATE WITH CADN AND PARTICIPATE
CC IN THE TRANSMISSION OF DEVELOPMENTAL INFORMATION. CAN ASSOCIATE
CC WITH ALPHA-CATENIN. CYTOPLASMIC ISOFORM ACCUMULATES THROUGH WG
CC SIGNALING. ARM FUNCTION IN WG SIGNAL TRANSDUCTION IS REQUIRED
CC EARLY IN DEVELOPMENT FOR DETERMINATION OF NEUROBLAST FATE. ARM AND
CC ABL PROTEINS FUNCTION COOPERATIVELY AT ADHERENS JUNCTIONS IN BOTH
CC THE CNS AND EPIDERMIS.
CC SUBCELLULAR LOCATION: CYTOPLASMIC AND MEMBRANE-ASSOCIATED WITH THE
CC INNER SURFACE OF CELL MEMBRANE.
CC ALTERNATIVE PRODUCTS: 2 ISOFORMS: CYTOPLASMIC (SHOWN HERE) AND
CC NEURAL; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC TISSUE SPECIFICITY: CYTOPLASMIC ISOFORM IS PREDOMINANT BEFORE GERM
CC BAND RETRACTION, AFTER RETRACTION AND DURING LARVAL STAGES, IT IS
CC FOUND IN HIGH LEVELS IN SPECIFIC CELLS ALONG THE CNS MIDLINE.
CC NEURAL ISOFORM IS FIRST SEEN AFTER GERM BAND RETRACTION IN THE
CC AXON TRACTS OF THE CNS, ALSO PRESENT IN AXONS DURING LARVAL STAGES
CC AND ACCUMULATES IN THE MOTOR NEURONS OF THE SEGMENTAL AND
CC INTERSEGMENTAL NERVES AS THEY EXIT THE CNS. BOTH ISOFORMS
CC ACCUMULATE IN THE PNS.
CC DEVELOPMENTAL STAGE: PRESENT AT ALL STAGES, BUT REACHED THE
CC HIGHEST LEVELS DURING EARLY TO MID-EMBRYOGENESIS.
CC PTM: PHOSPHORYLATED ON SER, THR AND TYR RESIDUES. LEVEL OF
CC PHOSPHORYLATION VARIES BOTH DURING EMBRYONIC DEVELOPMENT AND FROM
CC EMBRYONIC TISSUE TO TISSUE. SGG IS REQUIRED FOR PHOSPHORYLATION
CC AND WG SIGNAL NEGATIVELY REGULATES ARM PHOSPHORYLATION.
CC HYPOPHOSPHORYLATED FORM OF ARM INCREASES IN STEADY-STATE LEVELS.
CC SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC SIMILARITY: CONTAINS 12.5 ARM REPEATS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL; X54468; CAA38350.1;
CC EMBL; AF001213; AAB58731.1;
CC EMBL; AE003422; AAF45688.1; ALT_INIT.
CC EMBL; AL021106; CAA15946.1;
CC EMBL; AL021086; CAA15946.1; JOINED.
CC EMBL; AL021086; CAA15935.1;
CC EMBL; AL021106; CAA15935.1;
CC PIR; A31861; A31861.
CC HSSP; O02248; 1DOW.
CC TRANSFAC; T02977;
CC FlyBase; FBgn0000117; arm.
CC InterPro; IPR000225; Armadillo.
CC Pfam; PF00514; Armadillo_seg; 12.
CC SMART; SM00185; ARM; 11.
CC PROSITE; PS0176; ARM_REPEAT; 9.
CC Developmental protein; Segmentation polarity protein; Phosphorylation;
CC Cell adhesion; Cytoskeleton; Structural protein; Phosphorylation;
CC Cell adhesion; Cytoskeleton; Structural protein; Phosphorylation;
CC Cell adhesion; Cytoskeleton; Structural protein; Phosphorylation;

FT DOMAIN 1 158 ASP/GLU-RICH (ACIDIC).
FT REPEAT 159 200 ARM 1.
FT REPEAT 201 242 ARM 2.
FT REPEAT 243 284 ARM 3.
FT REPEAT 285 326 ARM 4.
FT REPEAT 327 368 ARM 5.
FT REPEAT 369 410 ARM 6.
FT REPEAT 411 449 ARM 7.
FT REPEAT 450 496 ARM 8.
FT REPEAT 497 538 ARM 9.
FT REPEAT 539 584 ARM 10.
FT REPEAT 585 608 ARM 11.
FT REPEAT 609 647 ARM 12.
FT REPEAT 648 689 ARM 13 (INCOMPLETE).
FT DOMAIN 690 843 ASP/GLU-RICH (ACIDIC).
FT VARSPLIC 718 843 LGPEAEVGLYGQPPSVHSSHGGRAPHOQGYDPLPDSMQ
GLEISSVGGGAGGAGNGAVGGGNGNIGAIPPSGA
PTSPYSMDVGEIDAGALNFDLAMPPTPDNNLNLAAYGD
TDC -> ILXO (IN NEURAL ISOFORM).
FT SEQUENCE 843 AA; 91152 MW; 40DAD6FB83163049 CRC64;

Query Match 80.9%; Score 195; DB 1; Length 843;
Best Local Similarity 78.3%; Pred. No. 3.7e-18;
Matches 36; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

OY 1 RHQAEQAQNAVRHLYGLPVVVKLLHPPSHWPLIKATVGLIRNLAL 46
II :::::IIIIIIIIII :::::IIIIIIIIII :::::IIIIIIIIII
DB 482 RHVDSQAQNAVRHLYGLSVIVKLLHPPSRWPLIKAVIGLIRNLAL 527

RESULT 10
PLAK_MOUSE STANDARD; PRT; 621 AA.
AC Q02257;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Junction plakoglobin (Desmoplakin III) (Fragment).
GN JUP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92376536; PubMed=1509266;
RA Butz S.; Steppert J.; Weissig H.; Kemler R.;
RT "Plakoglobin and beta-catenin: distinct but closely related.";
RL Science 257:1142-1144(1992).
RN [2]
RP REVISIONS TO 294 AND 296.
RA Butz S.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: COMMON JUNCTIONAL PLAQUE PROTEIN. THE MEMBRANE-
CC ASSOCIATED PLAQUES ARE ARCHITECTURAL ELEMENTS IN AN IMPORTANT
CC STRATEGIC POSITION TO INFLUENCE THE ARRANGEMENT AND FUNCTION
CC OF BOTH THE CYTOSKELETON AND THE CELLS WITHIN THE TISSUE. THE
CC PRESENCE OF PLAQUEGLUBIN IN BOTH THE DESMOSOMES AND IN THE
CC INTERMEDIATE JUNCTIONS SUGGESTS THAT IT PLAYS A CENTRAL ROLE IN
CC THE STRUCTURE AND FUNCTION OF SUBMEMBRANOUS PLAQUES.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC IN A SOLUBLE AND MEMBRANE-
CC ASSOCIATED FORM.
CC -1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -1- SIMILARITY: CONTAINS AT LEAST 9 ARM REPEATS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to license@isb-sib.ch.

CC EMBL: M90365; AAB02885.1;
CC PIR: S35092; S35092;
CC HSP: Q02248; IDOW.
CC MGD: MGI:96650; JUP.
CC InterPro: IPR000225; Armadillo.
CC Pfam: PF00514; Armadillo_seg; 11.
CC SMART: SM00185; ARM; 8.
CC PROSITE: PS00176; ARM_REPEAT; 9.
CC Cell adhesion; Cytoskeleton; Structural protein; Repeat.
FT NON_TER 1 1
FT REPEAT 8 47 ARM 1.
FT REPEAT 92 131 ARM 2.
FT REPEAT 134 173 ARM 3.
FT REPEAT 218 257 ARM 4.
FT REPEAT 259 296 ARM 5.
FT REPEAT 299 340 ARM 6.
FT REPEAT 346 386 ARM 7.
FT REPEAT 388 427 ARM 8.
FT REPEAT 450 489 ARM 9.
SQ SEQUENCE 621 AA; 68111 MW; 17CF444607422BAA CRC64;

Query Match 78.8%; Score 190; DB 1; Length 621;
Best Local Similarity 73.9%; Pred. No. 1.2e-17;
Matches 34; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

OY 1 RHQAEQAQNAVRHLYGLPVVVKLLHPPSHWPLIKATVGLIRNLAL 46
II :::::IIIIIIIIII :::::IIIIIIIIII :::::IIIIIIIIII
DB 341 RHPEAQAQNSVRLNYGIPAIKVLKLNQNPVLRKATIGLIRNLAL 386

RESULT 11
PLAK_HUMAN STANDARD; PRT; 743 AA.
AC P14923;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Junction plakoglobin (Desmoplakin III).
GN JUP-OR DP3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89264555; PubMed=2726765;
RA Franke W.; Goldschmidt M.D.; Zimbelmann R.; Mueller H.M.;
RA Schiller D.L.; Cowin P.;
RT "Molecular cloning and amino acid sequence of human plakoglobin, the
common junctional plaque protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:4027-4031(1989).
CC -1- FUNCTION: COMMON JUNCTIONAL PLAQUE PROTEIN. THE MEMBRANE-
CC ASSOCIATED PLAQUES ARE ARCHITECTURAL ELEMENTS IN AN IMPORTANT
CC STRATEGIC POSITION TO INFLUENCE THE ARRANGEMENT AND FUNCTION
CC OF BOTH THE CYTOSKELETON AND THE CELLS WITHIN THE TISSUE. THE
CC PRESENCE OF PLAQUEGLUBIN IN BOTH THE DESMOSOMES AND IN THE
CC INTERMEDIATE JUNCTIONS SUGGESTS THAT IT PLAYS A CENTRAL ROLE IN
CC THE STRUCTURE AND FUNCTION OF SUBMEMBRANOUS PLAQUES.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC IN A SOLUBLE AND MEMBRANE-
CC ASSOCIATED FORM.
CC -1- SIMILARITY: BELONGS TO THE BETA-CATENIN FAMILY.
CC -1- SIMILARITY: CONTAINS 8 ARM REPEATS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AL031325; CAA21824.1; -
DR HSP; 094766; IFGG.
DR FlyBase; FBgn026081; EG:EG0007.5.
KW Transferase; Glycoprotein; Transmembrane; Signal-anchor.
FT DOMAIN 1 11 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 12 29 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT DOMAIN 30 313 LUMENAL (POTENTIAL).
FT ACT_SITE 259 259 CATALYTIC BASE (BY SIMILARITY).
FT METAL 170 170 MANGANESE (BY SIMILARITY).
FT CARBOHYD 97 97 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 313 AA; 35805 MW; 0141DCD4CD778436 CRC64;

Query Match 23.4%; Score 56.5; DB 1; Length 313;
Best Local Similarity 34.0%; Pred. No. 3.2;
Matches 17; Conservative 7; Mismatches 11; Indels 15; Gaps 3;

QY 3 QEAEMAQNAVRLHYGLPVPVVKLLH-----PPSHWPLIK---ATVGLIRNL 44
Db 64 QKAELTNA-----PYSNRLSHLMLPLHLHWIIVDTNATTPLVRLN 106

RESULT 14

ID PURT_METJA STANDARD; PRT; 393 AA.
AC Q58881;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable phosphoribosylglycinamide formyltransferase 2 (EC 2.1.2.-)
DE (GART 2) (GAR transformylase 2) (5'-phosphoribosylglycinamide
DE transformylase 2) (Formate-dependent GAR transformylase).
GN PURT OR MJ1486.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; Pubmed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073(1996).

CC -1- FUNCTION: CATALYZES TWO REACTIONS: THE FIRST ONE IS THE PRODUCTION
CC OF BETA-FORMYL GLYCINAMIDE RIBONUCLEOTIDE (GAR) FROM FORMATE, ATP
CC AND BETA GAR. THE SECOND, A SIDE REACTION, IS THE PRODUCTION OF
CC ACETYL PHOSPHATE AND ADP FROM ACETATE AND ATP (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: FORMATE + ATP + 5'-PHOSPHO-RIBOSYLGLYCINAMIDE
CC = 5'-PHOSPHORIBOSYL-N-FORMYLGLYCINAMIDE + ADP + PYROPHOSPHATE (BY
CC SIMILARITY).

CC -1- PATHWAY: THIRD STEP (FIRST OF TWO TRANSFORMYLATION REACTIONS)

CC IN DE NOVO PURINE BIOSYNTHESIS. THIS IS AN ALTERNATIVE ENZYME TO
CC THE PURN GAR TRANSFORMYLASE (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE PURK / PURT FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U67589; AAB99496.1; -
DR TIGR; MJ1486;
DR InterPro; IPR003135; ATP-grasp.
DR Pfam; PF02222; ATP-grasp; 1.
KW Purine biosynthesis; Transferase; Complete proteome.
FT SEQUENCE 393 AA; 43553 MW; A47801A0277B6B87 CRC64;

Query Match 23.4%; Score 56.5; DB 1; Length 393;
Best Local Similarity 45.9%; Pred. No. 4;
Matches 17; Conservative 4; Mismatches 13; Indels 3; Gaps 2;

QY 3 QEAEMAQNAVRLHYGLPVPVVKLLHP-PSHWPLIKATV 38
Db 298 QEMSEFEIHRVAILGLPVPSTKLHPGASH--VIKAEI 332

RESULT 15

ID PETD_SYNP2 STANDARD; PRT; 160 AA.
AC P28057;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE Cytochrome B6-F complex subunit 4 (17 kDa polypeptide).
GN PETD.
OS Synecococcus sp. (strain PCC 7002) (Agmenellum quadruplicatum).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxID=32049;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93043038; Pubmed=1421151;
RA Brand S.N., Tan X., Widger W.R.;
RT "Cloning and sequencing of the petBD operon from the cyanobacterium
RT Synecococcus sp. PCC 7002.";
RL Plant Mol. Biol. 20:481-491(1992).
CC -1- FUNCTION: THIS POLYPEPTIDE OF UNKNOWN FUNCTION IS ONE OF THE
CC COMPONENTS OF THE CYTOCHROME B6-F COMPLEX.
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B6-F ARE: CYTOCHROME B6,
CC 17 kDa POLYPEPTIDE (PETD), CYTOCHROME F AND THE RIESKE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY. CORRESPONDS TO THE
CC CARBOXYL END OF MITOCHONDRIAL CYTOCHROME B.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X63049; CAA44775.1; -
DR PIR; S18124; S18124.
DR PIR; S26194; S26194.

DR InterPro; IPR000179; Cyt_b_b6.

DR Pfam; PF00032; cytochrome_b_c; 1.

DR PROSITE; PS00193; CYTOCHROME_B_QQ; 1.

KW Electron transport; Photosynthesis; Transmembrane.

SQ SEQUENCE 160 AA; 17537 MW; E0D8081AE4CAEB4 CRC64;

Query Match 23.0%; Score 55.5; DB 1; Length 160;
Best Local Similarity 31.4%; Pred. No. 2.1;
Matches 16; Conservative 5; Mismatches 11; Indels 19; Gaps 2;

QY 5 AEMAQNAVRLHYGLPVPVVKLLHPPSHWP-----LIKATVGLIRNLAL 46

Db 16 AKLAQNMGNHYGEPA-----WPNDILTFPICITAGTIGLTGUA 56

Tue Jul 30 08:15:33 2002

Search completed: July 29, 2002, 16:09:57
Job time: 762 sec

us-09-641-104a-12.rsp

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 29, 2002, 16:12:11 ; Search time 124.4 seconds
(without alignments)
63.969 Million cell updates/sec

Title: US-09-641-104A-12
Perfect score: 241
Sequence: 1 RHOEAMQNAVRLHYGLPV.....PPSHWPLIKATVGLIRNLAL 46

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: / 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_19.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phase.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_rvirus.*
 - 16: sp_bacteriap.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	241	100.0	781	13	042486
2	238	98.8	780	13	090424
3	233	96.7	781	11	090335
4	216	89.6	769	5	09NL44
5	216	89.6	773	5	076152
6	202	83.8	821	5	061229
7	190	78.8	729	13	09PVF7
8	190	78.8	745	4	Q15151
9	190	78.8	745	4	Q9BWC4
10	190	78.8	745	11	P70565
11	188	78.0	806	5	Q25100
12	78.5	32.6	678	5	Q44326
13	63.5	26.3	1181	5	Q9U308
14	60.5	25.1	107	2	Q9RH13
15	59.5	24.7	351	17	Q97W55
16	58	24.1	359	11	Q9D102

ALIGNMENTS

RESULT 1
ID 042486 PRELIMINARY; PRT; 781 AA.
AC 042486;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE BETA CATENIN.
GN CHBCAT.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1].
RP SEQUENCE FROM N.A.
RC STRAIN=WHITE LECHORN; TISSUE=DORSAL SKIN;
RX MEDLINE=97464068; PubMed=9322759;
RA Lu J., Chuong C.M., Widelitz R.B.;
RT "Isolation and characterization of chicken beta-catenin.";
RL Gene 196:201-207(1997).
DR EMBL; U82964; AAB80856.1; --
DR HSSP; P35222; IG3J
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 12.
DR SMART; SM00185; ARM; 11.
DR PROSITE; PS50176; ARM_REPEAT; 8.
SQ SEQUENCE 781 AA; 85438 MW; 6D205D9A4DBAC562 CRC64;

Query Match 100.0%; Score 241; DB 13; Length 781;
Best Local Similarity 100.0%; Pred. No. 9.7e-25;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1. RHOEAMQNAVRLHYGLPVVVVKLLHPSPHWPLIKATVGLIRNLAL 46
Db 474 RHOEAMQNAVRLHYGLPVVVVKLLHPSPHWPLIKATVGLIRNLAL 519

Q94b70 arabidopsis
Q9d294 mus musculus
Q9q8r8 myxoma viru
Q9v318 drosophila
Q97059 drosophila
Q43028 schizosacch
Q93xf3 zea mays (m
Q97061 drosophila
Q23928 eleocharis
Q23927 eleocharis
Q9pcy0 xylella fas
Q9hnp7 halobacteri
Q9fj16 arabidopsis
Q35812 rattus norv
Q9u3a8 caenorhabdi
P90935 caenorhabdi
P90936 caenorhabdi
Q91081 human immu
Q9de1 xenopus lae
Q9dfb2 xenopus lae
Q24801 entamoeba h
Q912c5 streptomyce
Q76283 trypanosoma
O22220 arabidopsis
Q66972 feline immu
Q06643 nitrosococc
Q9rail nitrosococc
Q9kky5 vibrio chol
Q92412 mus musculus

```

RESULT 2
Q90424 ID Q90424 PRELIMINARY; PRT; 780 AA.
AC Q90424;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE B-CATENIN.
GN CTNNB.
OS Brachydanio rerio (zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96122902; PubMed=8562427;
RT Kelly G.M., Erezylmaz D.F., Moon R.T.;
RT "Induction of a secondary embryonic axis in zebrafish occurs following
RL the overexpression of beta-catenin.";
RL Mech. Dev. 53:261-273(1995).
DR EMBL; U41081; AAC59732.1; -.
DR HSSP; P35222; 1G3J.
DR ZFIN; ZDB-GENE-980526-362; cttnb.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 12.
DR SMART; SM00185; ARM; 11.
DR PROSITE; PS50176; ARM_REPEAT; 8.
SQ SEQUENCE 780 AA; 85542 MW; D7A1FB80F94066DC CRC64;

Query Match 98.8%; Score 238; DB 13; Length 780;
Best Local Similarity 97.8%; Pred. No. 2.5e-24; Indels 0; Gaps 0;
Matches 45; Conservative 1; Mismatches 0;

Qy 1 RHQEAQAQNAVRHYGLPVVVVKKLHPPSHWPLIKATVGLIRNLAL 46
Db 473 RHQEAQAQNAVRHYGLPVVVVKKLHPPSHWPLIKATVGLIRNLAL 518
|||||
|||||

RESULT 3
Q9D335 ID Q9D335 PRELIMINARY; PRT; 781 AA.
AC Q9D335;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE ADULT MALE COLON CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,
DE CLONE:9030417H18, FULL INSERT SEQUENCE.
GN CATNB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=COLON;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staebli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

```

```

RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK018515; BAB31250.1; -.
DR HSSP; P35222; 1G3J.
DR MCD; MGI:88276; Catnb.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 12.
DR SMART; SM00185; ARM; 12.
DR PROSITE; PS50176; ARM_REPEAT; 7.
SQ SEQUENCE 781 AA; 85546 MW; 937538C3B5CD75D1 CRC64;

Query Match 96.7%; Score 233; DB 11; Length 781;
Best Local Similarity 95.7%; Pred. No. 1.2e-23; Indels 0; Gaps 0;
Matches 44; Conservative 0; Mismatches 2;

Qy 1 RHQEAQAQNAVRHYGLPVVVVKKLHPPSHWPLIKATVGLIRNLAL 46
Db 474 RHQEAQAQNAVRHYGLPVVVVKKLHPPSHWPLIKATVGLIRNLAL 519
|||||
|||||

RESULT 4
Q9NL44 ID Q9NL44 PRELIMINARY; PRT; 769 AA.
AC Q9NL44;
DT 01-OCT-2000 (TEMBLrel. 15, Created)
DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE BETA-CATENIN.
GN CIBETA-CATENIN.
OS Ciona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;
OC Cionidae; Ciona.
OX NCBI_TaxID=7719;
RN [1]
RP SEQUENCE FROM N.A.
RA Imai K., Takada N., Satoh N., Satou Y.;
RT "An essential role of beta-catenin in the endoderm specification of
RT ascidian embryo.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB031543; BAA92185.1; -.
DR HSSP; P35222; 1G3J.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 11.
DR SMART; SM00185; ARM; 11.
DR PROSITE; PS50176; ARM_REPEAT; 6.
SQ SEQUENCE 769 AA; 84703 MW; F61CC489B436E1BC CRC64;

Query Match 89.6%; Score 216; DB 5; Length 769;
Best Local Similarity 89.1%; Pred. No. 2.7e-21; Indels 0; Gaps 0;
Matches 41; Conservative 2; Mismatches 3;

Qy 1 RHQEAQAQNAVRHYGLPVVVVKKLHPPSHWPLIKATVGLIRNLAL 46
Db 464 RHQEAQAQNAVRHYGLPVVVVKKLHPPSHWPLIKATVGLIRNLAL 509
|||||
|||||

RESULT 5
O76152 ID O76152 PRELIMINARY; PRT; 773 AA.
AC O76152;
DT 01-NOV-1998 (TEMBLrel. 08, Created)
DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE BETA-CATENIN.
OS Ciona savignyi.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Phlebobranchia;
OC Cionidae; Ciona.
OX NCBI_TaxID=51511;

```

```

RN  SEQUENCE FROM N.A.
RX  MEDLINE=98443204; PubMed=9769178;
RA  Yoshida S., Marikawa Y., Satoh N.;
RT  "Regulation of the trunk-tail patterning in the ascidian embryo: a
RT  possible interaction of cascades between lithium/beta-catenin and
RT  localized maternal factor pem.";
RL  Dev. Biol. 202;264-279(1998).
DR  EMBL; AB012160; BAA32789.1; -.
DR  HSP; P35222; IG3J.
DR  InterPro; IPR000225; Armadillo.
DR  Pfam; PF00514; Armadillo_seg; 11.
DR  SMART; SM00185; ARM; 9.
DR  PROSITE; PS50176; ARM_REPEAT; 7.
SQ  SEQUENCE 773 AA; 85217 MW; C1340CF82AFEDAB CRC64;

Query Match      89.68; Score 216; DB 5; Length 773;
Best Local Similarity 89.18; Pred. No. 2.7e-21;
Matches 41; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy  1 RHOEAMQNAVRLHYGLPVVVKLLHPPSHWPLIKATVGLIRNLAL 46
Db  463 RHPDAEMQNAVRLHYGLPVVVKLLHPPSRWPLIKAVVGLIRNLAL 508

RESULT 6
O61229 PRELIMINARY; PRT; 821 AA.
AC O61229;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE BETA-CATENIN.
OS Lytechinus variegatus (Sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Temnopleuroidea; Toxopneustidae;
OC Lytechinus.
OX NCBI_TaxID=7654;
RN 1
RP SEQUENCE FROM N.A.
RX MEDLINE=98104237; PubMed=9441670;
RA Miller J.R., McClay D.R.;
RT "Changes in the pattern of adherens junction-associated beta-catenin
RT accompany morphogenesis in the sea urchin embryo.";
RL Dev. Biol. 192:310-322(1997).
DR EMBL; U34814; AAC06340.1; -.
DR HSP; Q02248; IDOW.
DR InterPro; IPR002086; Aldehyde_dehydr.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 12.
DR SMART; SM00185; ARM; 12.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
DR PROSITE; PS50176; ARM_REPEAT; 8.
SQ SEQUENCE 821 AA; 89558 MW; 71E21D562A99C5AD CRC64;

Query Match      83.88; Score 202; DB 5; Length 821;
Best Local Similarity 82.68; Pred. No. 2.5e-19;
Matches 38; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy  1 RHOEAMQNAVRLHYGLPVVVKLLHPPSHWPLIKATVGLIRNLAL 46
Db  491 RHPGEMQNTVRLNYGIPVIVKLLHPPSRWPLIKATVGLIRNLAL 536

RESULT 7
Q9PVF7 PRELIMINARY; PRT; 729 AA.
AC Q9PVF7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

```

```

DE CELL-ADHESION PROTEIN PLAKOGLOBIN.
GN JUP;
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN 1
RP SEQUENCE FROM N.A.
RX MEDLINE=93386700; PubMed=10456847;
RA Cerdá J., Reidenbach S., Pratzel S., Franke W.W.;
RT "Cadherin-catenin complexes during zebrafish oogenesis: heterotypic
RT junctions between oocytes and follicle cells.";
RL Biol. Reprod. 61:692-704(1999).
DR EMBL; AF099738; AAD56592.1; -.
DR HSP; Q02248; IDOW.
DR ZFIN; ZDB-GENE-991207-22; jup.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 11.
DR SMART; SM00185; ARM; 8.
DR PROSITE; PS50176; ARM_REPEAT; 8.
SQ SEQUENCE 729 AA; 80033 MW; 91E00417B4FD8CEE CRC64;

Query Match      78.88; Score 190; DB 13; Length 729;
Best Local Similarity 73.98; Pred. No. 1e-17;
Matches 34; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy  1 RHOEAMQNAVRLHYGLPVVVKLLHPPSHWPLIKATVGLIRNLAL 46
Db  454 RHEAETAQNAVRMHYGIPIVKKLNQYHWPVAVVGLIRNLAL 499

RESULT 8
Q15151 PRELIMINARY; PRT; 745 AA.
AC Q15151; Q15093;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PLAKOGLOBIN.
GN JUP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1
RP SEQUENCE FROM N.A.
RX MEDLINE=89264555; PubMed=2726765;
RA Franke W.W., Goldschmidt M.D., Zimbelmann R., Mueller H.M.,
RA Schiller D.L., Cowin P.;
RT "Molecular cloning and amino acid sequence of human plakoglobin, the
RT common junctional plaque protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:4027-4031(1989).
RN 2
RP SEQUENCE FROM N.A.
RA Zimbelmann R.;
RN Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN 3
RP SEQUENCE OF 239-409 FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=96157724; PubMed=8576101;
RA Ozawa M., Nukui K., Toyoyama H., Ohi Y.;
RT "Cloning of an alternative form of plakoglobin (gamma-catenin) lacking
RT the fourth armadillo repeat.";
RL J. Biochem. 118:836-840(1995).
RN 4
RP SEQUENCE FROM N.A.
RX PubMed=11016852;
RA Whittock N.V., Eady R.A.J., McGrath J.A.;
RT "Genomic Organization and amplification of the human plakoglobin
RT gene.";
RL Exp. Dermatol. 9:323-326(2000).

```

DR EMBL; Z68228; CAA92522.1; -
DR EMBL; D50808; BAA09435.1; -
DR EMBL; AF306723; AAG16727.1; -
DR EMBL; AF233882; AAG16727.1; JOINED.
DR HSSP; Q02248; 2BCT.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 11.
DR SMART; SM00185; ARM; 9.
DR PROSITE; PS00176; ARM_REPEAT; 7.
SQ SEQUENCE 745 AA; 81744 MW; 3519A0973748BCF4 CRC64;

Query Match 78.8%; Score 190; DB 4; Length 745;
Best Local Similarity 73.9%; Pred. No. 1e-17; 4; Indels 0; Gaps 0;
Matches 34; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 1 RHQEAEMQNAVRHLHYGLPVVVKLLHPPSHWPLIKATVGLIRNLAL 46
Db 465 RHPEAEMAQNSVRLNYGIPAIKLLNQNPQWPLVKATIGLIRNLAL 510

RESULT 9
Q9BWC4 PRELIMINARY; PRT; 745 AA.
ID Q9BWC4
AC Q9BWC4; (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE JUNCTION PLAKOGLOBIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG CARCINOMA;
RA Strausberg R.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA, CHORIOCARCINOMA;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC000441; AAH00441.1; -
DR EMBL; BC011865; AAH11865.1; -
DR HSSP; Q02248; 2BCT.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 11.
DR SMART; SM00185; ARM; 12.
DR PROSITE; PS00176; ARM_REPEAT; 7.
SQ SEQUENCE 745 AA; 81726 MW; 34DF7BFB4748BCF4 CRC64;

Query Match 78.8%; Score 190; DB 4; Length 745;
Best Local Similarity 73.9%; Pred. No. 1e-17; 4; Indels 0; Gaps 0;
Matches 34; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 1 RHQEAEMQNAVRHLHYGLPVVVKLLHPPSHWPLIKATVGLIRNLAL 46
Db 465 RHPEAEMAQNSVRLNYGIPAIKLLNQNPQWPLVKATIGLIRNLAL 510

RESULT 10
P70565 PRELIMINARY; PRT; 745 AA.
ID P70565
AC P70565;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PLAKOGLOBIN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=VAGINA;
RA Umekita Y., Liao S.;
RT "Molecular cloning and sequencing of the rat plakoglobin cDNA."
RL Submitted (AUG1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=VAGINA;
RA Hiipakka R.A.;
RL Submitted (MAY1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; F058858; AAB06317.1; -
DR HSSP; Q02248; 2BCT.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 11.
DR SMART; SM00185; ARM; 8.
DR PROSITE; PS00176; ARM_REPEAT; 7.
SQ SEQUENCE 745 AA; 81777 MW; 9E2F52910A7ACD41 CRC64;

Query Match 78.8%; Score 190; DB 11; Length 745;
Best Local Similarity 73.9%; Pred. No. 1e-17; 4; Indels 0; Gaps 0;
Matches 34; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 1 RHQEAEMQNAVRHLHYGLPVVVKLLHPPSHWPLIKATVGLIRNLAL 46
Db 465 RHPEAEMAQNSVRLNYGIPAIKLLNQNPQWPLVKATIGLIRNLAL 510

RESULT 11
Q25100 PRELIMINARY; PRT; 806 AA.
ID Q25100
AC Q25100;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE BETA-CATENIN.
OS Hydra magnipapillata (Hydra).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;
OC Hydridae; Hydra.
OX NCBI_TaxID=6085;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WILDTYPE 105;
RX MEDLINE=9625271; PubMed-8654977;
RA Hobmayer E., Hatta M., Fischer R., Fujisawa T., Holstein T.W.,
RA Sugiyama T.;
RT "Identification of a Hydra homologue of the beta-
catenin/plakoglobin/armadillo gene family."
RL Gene 172:155-159(1996).
DR EMBL; U36781; AAC47137.1; -
DR HSSP; Q02248; 2BCT.
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 10.
DR SMART; SM00185; ARM; 10.
DR PROSITE; PS00176; ARM_REPEAT; 7.
SQ SEQUENCE 806 AA; 90462 MW; 689E5E982CD5051A CRC64;

Query Match 78.0%; Score 188; DB 5; Length 806;
Best Local Similarity 73.9%; Pred. No. 2.1e-17; 6; Indels 0; Gaps 0;
Matches 34; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 RHQEAEMQNAVRHLHYGLPVVVKLLHPPSHWPLIKATVGLIRNLAL 46
Db 533 RHPEAEMAQNSVRLNYGIPAIKLLNQNPQWPLVKATIGLIRNLAL 578

RESULT 12
O44326 PRELIMINARY; PRT; 678 AA.
ID O44326
AC O44326;

[illegible]

RT -Sequence analysis of 55G; cosmid clone of *Zymomonas mobilis* 2M4
containing rRNA operon.
RL Submitted (AUG-1998); EMBL/GenBank/DBJ databases.
RR EMBL; AF088897; AAF18291.1; -
DR InterPro: IPR003844; UPF0060.
DR Pfam: PF02694; UPF0060; 1.
KW Hypothetical protein.
KW SEQUENCE 107 AA: 11921 MW: 80AD302AC5926272 CRC64:

Query Match	25.1%	Score 60.5;	DB 2;	Length 107;
Best-Local Similarity	34.1%	Pred. No. 0.96;		
Matches	15;	Conservative	7;	Mismatches 15; Indels 7; Gaps
QY	5	AEANQAVNLVHGLYLPVWVKLL-----HPSHWPLKATVGLI 41		
		: : : :		
DB	52	AENAKAYAVGGYIIMSLWSKWKVEATPPEHMDLTGAFAFLV 95		
RESULT	15.			
Q97W55	AC	PRELIMINARY;	PRT;	351 AA.
ID	Q97W55			
CD	Q97W55;			
DT	01-OCT-2001 (TREMBlrel. 18, Created)			
DT	01-OCT-2001 (TREMBlrel. 18, Last sequence update)			
DT	01-DEC-2001 (TREMBlrel. 19, Last annotation update)			
DE	HYPOTHETICAL GTP BINDING PROTEIN SSO2385.			
DE	SSO2385.			
GN	OS	Sulfolobus solfataricus.		
OC	Archaea; Crenarchaeota;	Sulfolobales; Sulfolobaceae; Sulfolobus.		
OX	NCBI_TaxID=2287;			
ON	[1];			
RP	SEQUENCE FROM N.A....			
RC	STRAIN-ATCC 35092 / DSM 1617 / P2;			
RX	MEDLINE=21332296; PubMed=11427726;			
RA	She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,			
RA	Awayez M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtis B.A.,			
RA	De Moors A., Erasus G., Fletcher C., Gordon P.M.K.,			
RA	Heikamp-de Jong I., Jeffries A.C., Kozaera C.J., Medina N., Peng X.,			
RA	Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,			
RA	Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,			
RA	Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;			
RT	"The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).			
DR	EMBL; AE006839; AAK42533.1; -			
DR	InterPro; IPR004095; TGS.			
DR	Pfam; PF02824; TGS; 1.			

KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 351 AA; 39916 MW; 046A96BF004865DE CRC64;

Query Match 24.7%; Score 59.5; DB 17; Length 351;
Best Local Similarity 54.2%; Pred. No. 4.8;
Matches 13; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

QY 20 VVKKLLRPPSHWPLIKATVGLIRN 43
| : : : : | | | : : : : |
Db 120 VOIQLVNPPKSLPLSK-TIGLIRN 142

Search completed: July 29, 2002, 16:12:12
Job time: 482 sec

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)